

AF052135 1462 bp mRNA linear PRI 05-AUG-1998

LOCUS AF052135

DEFINITION Homo sapiens clone 23625 mRNA sequence.

ACCESSION AF052135

VERSION AF052135.1 GI:3360444

KEYWORDS FLI CDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1462)

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

TITLE Andersson, B., Wentland, M.A., Ricafrente, J.Y., Liu, W. and Gibbs, R.A.

JOURNAL A 'double adaptor' method for improved shotgun library construction

MEDLINE Anal. Biochem. 236 (1), 107-113 (1996)

66207227

PubMed 8619474

REFERENCE 2 (bases 1 to 1462)

AUTHORS Yu, W., Andersson, B., Worley, K.C., Muzny, D.M., Ding, Y., Liu, W.,

TITLE Ricafrente, J.Y., Wentland, M.A., Lennon, G., and Gibbs, R.A.

JOURNAL Large-scale concatenation cDNA sequencing

MEDLINE Genome Res. 7 (4), 353-358 (1997)

97264341

PubMed 9110174

REFERENCE 3 (bases 1 to 1462)

AUTHORS Yu, W., Sarginson, J., and Gibbs, R.A.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-1998) Molecular and Human Genetics, Baylor

College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA

FEATURES

source

1..1462

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="I.M.A.G.E. Consortium clone ID 23625"

/sex="female"

/tissue_type="brain"

/clone_lib="INIB"

/dev_stage="infant"

ORIGIN

Alignment Scores:

Pred. No.: 3,57e-171 Length: 1462

Score: 2208.00 Matches: 424

Percent Similarity: 100.008 Conservative: 0

Best Local Similarity: 100.008 Mismatches: 0

Query Match: 100.008 Gaps: 0

DB: 9

US-09-831-452-1 (1-424) x AF052135 (1-1462)

CY 1 MetSerAspHisGlyAspValSerLeuProGluuAspArgValArgLeuSerGln 20

DB 117 ATGTCTGACCACTGGAGATGCGACCTCCGCCGAGACCGGGTGAAGCTCTCTCCAG 176

CY 21 LeuGlySerAlaValGluValaGluuAspGluLeuProArgArgGlyPheArgSerGly 40

DB 177 CTGGGTGTGGTGGTGAAGTGAAGACATCCACCCCTGGTACTTCCGCTCTGCA 236

CY 41 ValGluLeuLeuArgMetAlaSerLeuLeuSerGluGluGluValaGluLeuAlaPhe 60

DB 237 GTTGAGATTATCCGATGCGATTCATTCCTCGAGAGAGGCAACATTGAATCCCTTC 296

CY 61 TLeuLeuValAsnLysTyrIleThrLeuPheIleGluLeuLeuProLysHisArgAspTyr 80

DB 297 ATCTCTATTAACAAGATATACAGCTCTTATTAAGACAACTACCAAAATACGAGATTAC 356

CY 81 LysSerAlaValIleProGluLysAspThrValLysLysLeuLysGluLeuAlaPhe 100

DB 357 AATTCCTCTCTATTCCTGAAAGAAAGACAGTAAAGAAATTAAAGAGATTGATTT 416

CY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120

DB 417 CCCAAAGCAAGAGCTGAAGGACAGAGCTGTAAACGATATACCAAAATATATACAGA 476

CY 121 TYAANGGUGUUGLYSLYSLYSGUUAAGUGUUAUUAARGANMETALALEGUGU 140

DB 477 TATTAATGAAGAAAGAAAGAGAGAGAGAGATTTGGCCCGGAAACATGGCCATCCAGCAA 536

CY 141 GULUUGUGUUGGUGUUGGUGUUGUUAUUAAGUGUUGUUGUUGUUGUUGU 160

DB 537 GAGCTGGAAAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 536

CY 161 GINPHEHIALAPHEGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 180

DB 597 CAGTTCATGCTCTTGAAGAGAGATGATCCGAAACAGAGAGAGAGAGAGAGAG 656

CY 181 TLEVALGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGU 200

DB 657 ATTGTACAGAGATTGGAGAGATGATGATGATGATGATGATGATGATGATGAT 716

CY 201 GULYSPROSERLEUASPVALPHEPROTHREUTHYALSERSERILEGINPROSERASP 220

DB 717 GAGAGCCCTCTTGAAGTGTCTCCACCTTAACAGCTCATTCATACAGCCTTCAGAC 776

CY 221 CYHISTHTRVALXPROALALYSPROVALVALASPARSERLEUASPARG 240

DB 777 TGTACACACTGTAAAGGCGAGTAAAGCCTGATGATGATGATGATGATGATGAT 836

CY 241 ALALEUSERSNSERGLUSERILEPROTHRIEASPGILEUARGHISVALVALPRO 260

DB 837 GCACCTAGCACTCAAGAAAGATTTCCCATTCATGATGATGATGATGATGATGAT 896

CY 261 GLYARGLEUCYSPROGLPHELEUGUUAUUAASERIALAANTRAALARGLYVALGLU 280

DB 897 GGGCGGCTGTCACAGATTCTCCAGTACAGTGCACACACTGCGGGGAGTGGAG 956

CY 281 THRCSGLYLELEUCYSGLYLYSLYSLYSLYSLYSLYSLYSLYSLYSLYSLY 300

DB 957 ACATGGAGATCTCTGTGAGAAATGATGATGATGATGATGATGATGATGATGAT 1016

CY 301 TLEPROLYGINSERIALGLYSERAPTYCYSEANTHRGUAANGUGUUGUUGUUG 320

DB 1017 ATCCCAAGCAAGGTGCTGGGTCTATTAATGCAACAGAGAGAGAGAGAGAGAG 1076

CY 321 LEULLEGNASPGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 340

DB 1077 CTCATACAGAGATCAAGAGGCTCATACAGTGGGCTGATTCATCAACCCACAG 1136

CY 341 THRALAPHELEUSERSERVALASPLEHISTHRSYCSERTYRGIMETLEUPRO 360

DB 1137 ACCGCTTCTCTCCAGTGTCACTCACTCACTCACTCACTCACTCACTCACTCA 1196

CY 361 GLUSERVALALALEVALCYSESPROLYSPHEGUGUUGUUGUUGUUGUUGUUG 380

DB 1197 GAGTACAGTGCATTTGTTGCTCCCAAGTTCAGAGAACTGAGATCTTTAACTACT 1256

CY 381 ASPHISGLYLEUGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 400

DB 1257 GACCATGAGCTGAGAGAGATTTCTCTGTCGCAAGAAAGATTTCAACACAGAGAG 1316

CY 401 ASPPROLEUPHECYSESPROLYSPHEGUGUUGUUGUUGUUGUUGUUGUUG 420

DB 1317 GATCCACCTCTGTTCTGATGAGTGCAGCAGGAGTGTGTGAGACAGAGAGTATC 1376

CY 421 THRASPLEUARG 424

DB 1377 ACAGACCTTCA 1388

RESULT 2

E35541 1910 bp DNA linear PAT 31-JAN-2002

LOCUS E35541

DEFINITION Protein AMSH and CDNA thereof.

ACCESSION E35541

VERSION E35541.1 GI:18624562

KEYWORDS UP 2000139469-A/1.

SOURCE Homo sapiens (human)

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 23, 2004, 16:45:15 ; Search time 5360 Seconds

(without alignments)
3428.628 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208

Sequence: 1 MSBHGVSLEPDEBRVRLSQ.....FCGSHVTYVDAVITIDLR 424

Scoring table:

BIOSUX62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgm2.1/USPTO.spool/us09831452/runat.17082004.151159.23641/app_query.fasta_1.563
-DB=Cgm2b1 -QMT=fastcap -SUFFIX=p2n.rge -MTMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -STAR=1 -END=-1 -MATRIX=Diosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09831452 @CGN 1.1 3731 @runat.17082004.151159.23641 -NCPU=6 -ICPU=3
-NO MAP -JARGOUTERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DEDEXT=7

Database :

GenEmbl:*
1: gb_hg:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pun:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of residues predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2208	100.0	1462	9	AF052135	AF052135 Homo sapi
2	2208	100.0	1910	6	E35541	E35541 Protein AMS
3	2208	100.0	1930	9	HSU73582	U73522 Homo sapien
4	2208	100.0	1933	9	BC007682	BC007682 Homo sapi
5	2208	100.0	2003	6	AX100575	AX100575 Sequence
6	1913	86.6	1544	10	AY083159	AY083159 Rattus no
7	1913	86.6	1735	10	BC061711	BC061711 Rattus no
8	1904	86.2	1384	6	E35542	E35542 Protein AMS
9	1904	86.2	2094	10	BC025111	BC025111 Mus muscu
10	1904	86.2	2104	10	AB010123	AB010123 Mus muscu
11	1904	86.2	2122	10	BC003497	BC003497 Mus muscu
12	1904	86.2	2172	10	BC006939	BC006939 Mus muscu
13	1409	63.8	2455	5	BC055512	BC055512 Danio rer
14	1404	63.6	2457	5	AY398309	AY398309 Danio rer
15	1269.5	57.5	1601	9	AB010120	AB010120 Homo sapi
16	1269.5	57.5	1621	6	BD186116	BD186116 Apoptosis
17	1269.5	57.5	1621	6	BD187452	BD187452 Apoptosis
18	1269.5	57.5	2006	9	AK056086	AK056086 Homo sapi
19	1269.5	57.5	2010	9	BC010846	BC010846 Homo sapi
20	1263.5	57.2	1717	10	BC018343	BC018343 Mus muscu
21	1262.5	57.2	1392	6	BD186118	BD186118 Apoptosis
22	1262.5	57.2	1392	6	BD187454	BD187454 Apoptosis
23	1262.5	57.2	1428	6	BD186117	BD186117 Apoptosis
24	1262.5	57.2	1428	6	BD187453	BD187453 Apoptosis
25	1262.5	57.2	1690	10	AB066213	AB066213 Mus muscu
26	1262.5	57.2	2008	10	AB066211	AB066211 Mus muscu
27	1262.5	57.2	2206	9	AK092946	AK092946 Homo sapi
28	1257	56.9	1828	5	BC061390	BC061390 Silurana
29	1237	56.0	10520	9	HSM809247	BX64906 Homo sapi
30	1236	55.7	1973	6	AX082297	AX082297 Sequence
31	1229.5	55.7	4052	9	AB037794	AB037794 Homo sapi
32	1105	50.0	1794	5	AF164597	AF164597 Lapevis h
33	1092	49.5	1964	5	BC046033	BC046033 Danio rer
34	871.5	39.5	1717	3	AY075309	AY075309 Drosophila
35	865	39.2	1774	3	AK114665	AK114665 Clona int
36	857.5	38.8	164603	2	AC119466	AC119466 Rattus no
37	857.5	38.8	223966	2	AC094504	AC094504 Rattus no
38	815.5	36.9	1898	6	AX746569	AX746569 Sequence
39	815.5	36.9	1898	9	AK090678	AK090678 Homo sapi
40	806	36.5	1356	10	AB066212	AB066212 Mus muscu
41	777.5	35.2	1246	6	AX780672	AX780672 Sequence
42	777.5	35.2	1246	6	AX780673	AX780673 Sequence
43	758	34.3	1041	6	BD264633	BD264633 Compositi
44	758	34.3	1041	6	AR238036	AR238036 Sequence
45	758	34.3	1041	6	AR257577	AR257577 Sequence

RESULT 1

ALIGNMENTS

AF052135 1462 bp mRNA linear PRI 05-AUG-1998

LOCUS AF052135 1462 bp mRNA linear PRI 05-AUG-1998

DEFINITION Homo sapiens clone 23625 mRNA sequence.

ACCESSION AF052135

VERSION AF052135.1 GI:3360444

KEYWORDS FLI CDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1462)

AUTHORS Anderson,B., Wentland,M.A., Ricafrente,J.Y., Liu,W. and Gibbs,R.A.

TITLE A 'double adaptor' method for improved shotgun library construction

JOURNAL Anal. Biochem. 236 (1), 107-113 (1996)

PMID 86207227

PMED 8619474

REFERENCE 2 (bases 1 to 1462)

AUTHORS Yu,W., Anderson,B., Worley,K.C., Muzny,D.M., Ding,Y., Liu,W., Ricafrente,J.Y., Wentland,M.A., Lennon,G. and Gibbs,R.A.

TITLE Large-scale concatenation cDNA sequencing

JOURNAL Genome Res. 7 (4), 353-358 (1997)

PMID 9110174

PMED 9110174

REFERENCE 3 (bases 1 to 1462)

AUTHORS Yu,W., Sargison,J. and Gibbs,R.A.

TITLE Direct Submission

JOURNAL Submitted (05-MAR-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA

FEATURES

source

1. 1462

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="I.M.A.G.E. Consortium clone ID 23625"

/sex="female"

/tissue_type="brain"

/clone_lib="INIB"

/dev_stage="infant"

ORIGIN

Alignment Scores:

Pred. NO: 3.57e-171 Length: 1462

Score: 2208.00 Matches: 424

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-09-831-452-1 (1-424) x AF052135 (1-1462)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20

DB 117 ATGCTGACCTGAGAGATGAGCTCCGCCGAGAACCGGGGCTCTCTCCAG 176

QY 21 LeuGlySerAlaValGluValAsnGluAspIleProGArgArgTyrPheArgSerGly 40

DB 177 CTGGGTAGTGGCGTAGAGGTGAATGAAGACATTCACCCGTCGGTACTTCCGCTCGGA 226

QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluValAsnIleGluHisAlaPhe 60

DB 237 GTTAGAGATTATCCGAATGGCATTCATTACTGAGAGAGCAACATTTGACATGCTTC 296

QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80

DB 297 ATCTCTATTAACAAGTATATCACTCTTATTGAGAACTACCAAAACATCGAGATTAC 356

QY 81 LysSerAlaValIleProGluLysAspThrValLysLysLeuGluIleAlaPhe 100

DB 357 AAATCTGCTGCATCTCCGAAAGAAAGACACAGTAAAGAAATTAAAGGATTGCAATT 416

QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGln 120

DB 417 CCAAGAGAGAGAGCTGAGAGAGAGCTGTTAAACGATATACCAAGATATACAGAA 476

QY 121 TYRAsnGluGluLysLysLysGluValGluGluLeuValAlaArgAsnMetAlaIleGlnGln 140

DB 477 TTTATATGAAGAAAGAAAGAAAGAAAGAGAGAAATTGGCCCGGAAACATGGCCATCCAGAA 536

QY 141 GluLeuGluLysGluLysGluArgValAlaGlnGlnLysGlnGlnGluGlnGlu 160

DB 537 GAGCTGGAAAGAAAGAAAGAAAGAGAGGTAGCACAAAGAGCGAGCATTTGGAAACAGAA 596

QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluLys 180

DB 597 GATTTCATGCTTTCGAGAGATGATCCGGAACCGAGAGCTAGAAAAAGCCACTGAA 656

QY 181 IleValGlnGluPheGluLysValAspProGluLeuGluGlyProLeuValProAspLeu 200

DB 657 ATTGTACAGAGATTGGGAAGTAGACCTGCTAGGTGGCCGCTAGTCCGACTTG 716

QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220

DB 717 GAGAAAGCCCTCTTAGATGTGTCTCCCACTTAACAGTCTATCCATACAGCTTCAGAC 776

QY 221 CysHisThrThrValArgProAlaLysProValAlaAspArgSerLeuLysProGly 240

DB 777 TGTCAACAAGCTGTAGAGCCAGTAAAGCCACTGCTGGTGAGACAGTCTTGAACCTGGA 836

QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGluLeuArgHisValValPro 260

DB 837 GCACGAGCAACTCAGAAAGATTTCCCAATCAATGATGATGCTGATGCTGATGCT 896

QY 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaSerThrAlaArgGlyValGlu 280

DB 897 GGGCGGCTGCCCCACATTTCTCCAGTTACAGCTGCAACATCGCCCGGAGGTGAG 956

QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300

DB 957 ACATGTGAATTTCTCTGTGAAACCTGATGAGATGAATTTACATTACCAATGATCTC 1016

QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320

DB 1017 ATCCCAAGCAAGAGCTGGGTGATTTATTCGCAACAGAGAAAGAACTTTTC 1076

QY 321 LeuIleGlnAspGlnGlnGluLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340

DB 1077 CTCATACAGATACACAGAGGCTCATCACATCGGCTGATTTACTACACCCACAG 1136

QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro 360

DB 1137 ACCGGGTTTCTCCAGTGTGACCTACACACTGCTCTTACCAATATATGTTGCA 1196

QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380

DB 1197 GAGTCAGTAGCATTTGTTGCTCCCCCAAGTTCCAGAACTGATTTTAAACTACT 1256

QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisPheHisSerLys 400

DB 1257 GACCAATGACATAGAGAAATTTCTCTGTGCGCAAGAAAGATTTCACTCACAGCAAG 1316

QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420

DB 1317 GATCCACTCTCTTGTGTAGCTGACGACGACGACTGTGTGACAGACAGTACCATC 1376

QY 421 ThrAspLeuArg 424

DB 1377 ACAGACCTTCCA 1388

RESULT 2

E35541 1910 bp DNA linear PAT 31-JAN-2002

LOCUS E35541

DEFINITION Protein AMSH and cDNA thereof.

ACCESSION E35541

VERSION E35541.1 GI:18624562

KEYWORDS JP 2000139469-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1910)
AUTHORS Sugamura, K. and Tanaka, N.
TITLE Protein AMSH and CDNA thereof
JOURNAL Patent: JP 2000139469-A 1 23-MAY-2000;
SCIENCE & TECH AGENCY
COMMENT OS Homo sapiens (human)
PN JP 2000139469-A/1
PD 23-MAY-2000
PF 12-NOV-1998 JP 1998322674
PR

FEATURES
Source Location/Qualifiers
1..1910
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 5.05e-171 Length: 1910
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-831-452-1 (1-424) x B35541 (1-1910)

QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
DB 11 ATGCTTACCATGAGATGAGAGCTCCGCCGGAAGACGGGGAGAGGCTCTCTCCAG 70
QY 21 LeuGlySerAlaValGluValAsnGluAspLeuProProArgArgTyrPheArgSerGly 40
DB 71 CTGGGATGAGGGGAGAGGAGTGAAGACATTCACCCCTGGATCTTCCCTCGGA 130
QY 41 ValGluLeuLeuArgMetAlaSerIleTyrSerGluGluGlnIleGluHisAlaPhe 60
DB 131 GTTAGATATATCCGAATGGATCCATTACTCTGAGGAGGCAACATTTGAACATGCTTC 190
QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
DB 191 ATCTCTATATACAGATATATCAGCTCTTTATGAGAACTACCAAAACATCCAGATTAC 250
QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100
DB 251 AAATTCGCTCTCATTCCTGAAAAGAAAGACACAGTAAAGAAATTAAGAGATTGATTT 310
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGln 120
DB 311 CCCAAACAGAAAGAGCTGAGAGGAGAGCTTTAAACAGATATCCAAAGATATACGAA 370
QY 121 TyrAsnGluGluLysLysGluLysAlaGluLeuLeuAlaArgMetAlaIleGlnGln 140
DB 371 TATATATAGAAAAGAAAGAGAGAGAGCATTTGGCCGGAACATGGCCATCCAGAA 430
QY 141 GluLeuGluLysGluLysGluLysArgValAlaGlnGlnLysGlnGlnGlnGlnGln 160
DB 431 GAGCTGAAAAGGAAAAGAGAGGGGTGACACACAGAGAGAGCATTTGGAAACGAA 490
QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgLeuLys 180
DB 491 CAGTTCCATGCTTCGAGAGATGATCCGACACAGAGAGCTAAGAAAAGAGAGCTGAA 550

QY 181 IleValGlnGluPheGluLysValAspProGluLysGluLysProLeuValProAspLeu 200
DB 551 ATGTACAGAGAGTTGGAAAGAGAGACCTGGCTAGGTGGCCGCTAGTGGCTTGG 610
QY 201 GlnLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
DB 611 GAGAGCCCTCTCTTATGATGTGTCCACCTTAACAGTCTCATACAGCTTACAGAC 670
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
DB 671 TGTCACACACTGTAAAGCCAGCTTAAGCCACTGTGTGTGACAGGCTTGAAACCTGGA 730
QY 241 AlLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
DB 731 GCACAGAGCACTCAGAAAGTATTCACACATGATGATGGCCATGTGGTGGTGGCT 790
QY 261 GlyArgLeuCyseProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
DB 791 GGGCGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTGCACACACTGCGGGGAGTGGAG 850
QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
DB 851 ACATGTGGAATCTCTGTGGAAGAACTGAGAGAGATGAATTAACATTACCATGTCTC 910
QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
DB 911 ATCCCAACAAAGTGTGGTGTCTGATCTACTCAACACAGAGAGAGAAACCTTTTC 970
QY 321 LeuIleGlnAspGlnGlnGlnLysLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
DB 971 CTCATACAGAGATCAGAGAGGCTCTCATCACCTGGGTGATTTCATATCAACCCACACAG 1030
QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360
DB 1031 ACCGCGTCTCTCCAGTGTGACCTACCTACCTCACTGCTCTTACCAAGATGATGGCA 1090
QY 361 GluSerValAlaIleValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuThr 380
DB 1091 GAGTCAGTAGCCATGTTGTGCTCCCAAGTTCAGAAACGATCTTTAAACTACT 1150
QY 381 AspHisGlyLeuGlnGluLysSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
DB 1151 GACCATGAGCTAGAGAGATTTCTTCCTGTGGCCGAAAGATTTCAATCCACACAGCAAG 1210
QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
DB 1211 GATCCACCTCTGTCTGTAGCTGACGACGACGACTGTGTGACAGAGAGTGCATTC 1270
QY 421 ThrAspLeuArg 424
DB 1271 ACAGACCTTCCA 1282

RESULT 3
HSU73522
LOCUS HSU73522 1930 bp mRNA linear PRI 29-JUN-1999
DEFINITION Homo sapiens AMSH mRNA, complete cds.
ACCESSION U73522
VERSION U73522.1 GI:4098123
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1930)
AUTHORS Tanaka, N., Kaneko, K., Asao, H., Kasai, H., Endo, Y., Fujita, T.,
Takeshita, T. and Sugamura, K.
TITLE Possible involvement of a novel STR-associated molecule 'AMSH' in
intracellular signal transduction mediated by cytokines
JOURNAL J. Biol. Chem. 274 (27), 19129-19135 (1999)
MEDLINE 99315854
PUBMED 10383417
REFERENCE 2 (bases 1 to 1930)

Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, K., Buecwo, R.H., Schaefer, C.F., Bat, N.K., Hopkins, K.F., Jordan, H., Moore, T., Max, S., Wang, J., Hsieh, F., Diatchenko, L., Martusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stancovski, I., Soares, M.B., Bonaldi, M.F., Casavault, T.L., Schaefer, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S., Carrance, P., Prange, C., Rana, S., Loquellano, N.A., Peters, G.J., Adams, E., Mullany, S.J., Bosak, S., McMan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulik,S.W.,
Vallation,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Sachdev,A., Whiting,M., Kettelman,M., Madan,A., Rodrigues,S.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1933)
Strasberg,R.
Direct Submission
Submitted (11-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigr.nih.gov
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,B., Kwong,P., Latic,P., Legaspi,R.,
Mazuro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McGowell,J., Pearson,R., Startripop,S., Thomas,P.J., Touchman,J.W.,
Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAL Plate: 12 Row: 1 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1778303.
Location/Qualifiers
1. .1933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:3546 IMAGE:361837"
/tissue="Lymph, Burkitt lymphoma"
/clone_id="Nih_MGC_8"
/lab_host="DH10B-R"
/note="Vector: pCIB7"
1. .1933
/gene="PAMSH"
/db_xref="LocusID:10617"
/db_xref="MIM:606247"
17. .1291
/codon_start=1
/product="associated molecule with the SH3 domain of STAY"
/protein_id="AAH07682.1"
/db_xref="GI:14043382"
/translation="MSDHGVSLPPEDRVRLAQLGSAVEVNDIPRRYRSGVELI
RMAISYSEGNIEHAFILNKYITLFEKLPKRDYSAVYPEKKTIVKLIKIPAK
AEELAEILKRYTEYENBEKEAEELANNAIQDELEKQVAAQOQOQLEB
QFPAEEMIRNQELEKRIKIVGEFGVDPGLGGPLVDELEKSLDVPLTVSIEP
SDCHTVPAKPVDRSLKRGALNSSEITPDIGLHVVPGLICPPLISANNA
RGVETGGLCKMLNERTITRHUIPKOSASDVCANENEELFLIDQOGLITLGGI
HTHPTQAFSSVDLTHCSYQMLTPSVAIVSPKQSTGFFKLTIDGLEEISCRQ
KGFPHSKDPLFCSHIVTVVDRATVITIDR"

misc_feature 770. .1099
/note="Mov33; Region: Mov34/MPN/PAD-1 family. Members of
this family are found in proteasome regulatory subunits,
eukaryotic initiation factor 3 (eif3) subunits and
regulators of transcription factors. This family is also
known as the MPN domain and PAD-1-like domain. It has been
shown that this domain occurs in prokaryotes"
/db_xref="CDD:pfam01398"
ORIGIN
Alignment Scores:
Pred. No.: 5.13e-171 Length: 1933
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-831-452-1 (1-424) x BC007682 (1-1933)
Qy 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
Db 17 ATGTCGACCGACGAGATGTGACCTCCCGCCGAAAGACGGGTGAGGCTCTCTCCAG 76
Qy 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
Db 77 CTGGGTAGTCGGGTAGCGTAAATGACATTCACCCCGCTGGTACTCCGCTCTGGA 136
Qy 41 ValGluIleIleArgMetIleSerIleTyrSerGluGluGlyAsnIleGluIleAlaPhe 60
Db 137 GTTGAATTAATTCGAAATGATCATTCCTCTGAGAAAGGACATTAACATGCTTC 196
Qy 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluTyrLeuProIleHisIleArgAspTyr 80
Db 197 ATCCTTAATACAGATATACAGCTCTTATTTAGAAACCTACAAACATCGAGATTAC 256
Qy 81 LysSerAlaValIleProGluTyrLysAspThrValLysLeuLysGluIleAlaPhe 100
Db 257 AAATCGCTGCTACTTCTGAAAGAAAGACAGATTAAGAAATTAAGAGATTCATT 316
Qy 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrTrpGlu 120
Db 317 CCCAAAGCAGACGAGTGAAGCGAGCTGTTAACGATATACAAAGATATACAGA 376
Qy 121 TyrAsnGluGluLysLysLysGluAlaGluGluLeuAlaArgAsnMetAlaIleGln 140
Db 377 TATTAATGAAGAAAGAAAGAGAGACGAGATTCGCCGAAACATGCTCCAGCAA 436
Qy 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnLeuGlnGln 160
Db 437 GACCTGAAAGGAAAGAAAGAGAGGTAGCAACAGACAGACGACCAATTGGAAACGGA 496
Qy 161 GlnPheIleAlaPheGluLysIleArgAsnGlnGluLeuGluLysGluArgLeuLys 180
Db 497 CAGTTCCATGCTTCAGAGAGATGATCCGAAACAGAGACCTGAAAAAGAGCACTGAA 556
Qy 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 557 ATTGTACAGAGATTGGAGAGTAGACCTGTGCTTAGGTGGCCCTGATGCTGACTTG 616
Qy 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 617 GAAAGCCCTCTCTAATGTGTTCCCACTTAACAGTCTATTCATACAGCTTCAGAC 676
Qy 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
Db 677 TGTCACACACACTGTAGAGGCAAGTAAAGCCATGTGGTGAACAGGCTTGAAACCTGGA 736
Qy 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 737 GCACTAGACACTCAGAAAGTATTCACAAATCGATGATTCGCGCATGTGGTGGCT 796
Qy 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280

Db 797 GGGGGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTCCACAGCTGCCGGGAGTGGAG 856
Oy 281 Thrcysglylleuvcysglylylsleumetargansgluphetrhllethrhlsvallleu 300
Db 857 ACATGTGAATTCCTCTGTGMAAACTGATGAGGATGAATTTCCATTACCATGTTCTC 916
Oy 301 lleprolysglnseralagliserasptrycysanthrhlgluasnugluuglualeupe 320
Db 917 ATCCCAAGCAAGAGTCTGGGTCTGATTACTGCAACAGAGAGAGAGAACTTTTC 976
Oy 321 leuileglnaspnglnglnglyleuilethrleuglyltpillehsthrhlspthrhlgin 340
Db 977 CTCTACAGAGATCAGCAGGCGCTCATCAGCTGGGCGGATTCATCTCACCACACAG 1036
Oy 341 Thrhlalpheleusersevalaephleuilethrhlscysseertryglnmetleupepro 360
Db 1037 ACCCGTTCTCTCCAGTGTGACCTTCACTCTCTTACAGATGAGTGTGCCA 1096
Oy 361 Gluisevalalalevalcysseerprolyspheglnlunthrhlgluphepelysleuthr 380
Db 1097 GAGTCAGTACGATGTTGTTGCTCCCAAGTCCAGAACTGATCTTAAACTTAAT 1156
Oy 381 Asphlsglyleugluugluileseertrycysarglnlueglyphehlspthrhlseerly 400
Db 1157 GACCATGAGACTAGAGGAGATTTCTCTGTGCGCAAGAGATTTCACTCCACAGCAG 1216
Oy 401 AspProleupheCysseerCysseerhlsvallThrvalalaspargalavalThrle 420
Db 1217 GATCAGCTCTGTTCTCTTACCTGAGCCAGCTGAGTCTTTGGACAGAGAGAGCAGC 1276
Oy 421 Thraspheudarg 424
Db 1277 ACAGACCTTCGA 1288
RESULT 5
AX100575
LOCUS AX100575 2003 bp DNA linear PART 10-APR-2001
DEFINITION Sequence 1 from Patent WO0121794.
ACCESSION AX100575
VERSION AX100575.1 GI:13619578
KEYWORDS
SOURCE
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Itoh, F., Itoh, S., Heldin, C.H. and Ten-Dijke, P.N.
AUTHORS Smad associating polypeptides
TITLE Patent: WO 0121794-A 1 29-MAR-2001;
JOURNAL LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
Source
1. .2003
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
63_1337
/note="unamed protein product"
CDS
/codon_start=1
/protein_id="CAG36455.1"
/db_xref="GI:13619579"
/translation="MSDGDVSLPPEDRVVALSQLSAVEVNEDIPPRRVPSGVEII
RMSIYSEGNIEHAFILYKNTYITLIEKLPKRDYKSAVIEPKDITVKKIEIAPK
AEEKAKELLKYTEKTEYVNEKKKEEELARPMALIOELKXKOPVAOQKQLEOF
QFAHEEMIKKQELERKLTIVDFGVDFDGLGSPVLDLEKSLDFVFTLVYSIQP
SDCHTIVPAKPKPVVDRLKPGALSNSESLPTIDGLRHVVPRGLCPQPLQASANTA
RGVETGILCGKLMRNFITTHVILPQSGSDYCTNEEELFLIDQDGLITLGM
HTHPTQAFSLSDVLTHTGSCYQMLPEFSAIVLSPFOETGPFKLTNDHGLEISCRQ
KGFHPSKDPFLFCSSHVTVDRAVTIDLR"

Pred. No.: 5,386-171 Length: 2003
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-831-452-1 (1-424) x AX100575 (1-2003)
Oy 1 Metserasphtlsglyaspvalserleuproprogluaspargalargalaleusergln 20
Db 63 ATGTCTGACATGAGAGATGAGCTCCGCCCAAGACCGGATGAGGCTCTCCAG 122
Oy 21 leuglyseralavalgluvalasngluaspilueproproargargtrypheargsergly 40
Db 123 CTGGGTAGTGGGATGAGAGTGAATGAGACATTCACCCCTCGGTACTTCCCTGGA 182
Oy 41 Valgluilelleargmeralaserletryserglugluysanilegluhalphe 60
Db 183 GTTGAGATTATCCGATGACATTCATTACTCTGAGAGGACCATTTGAACATGCCCTTC 242
Oy 61 lleleutryasnlystryllethrleupheillegluysleuprolyshlsargasptry 80
Db 243 ATCTCTTAACAAGTATATCACGCTCTTTATGGAACATACCAAAACATCGAGATTAC 302
Oy 81 lysseralavalilueprogluyleaspthrvalylslyleuylgluilealphe 100
Db 303 AATCTGCTGATCTCTCGAAAGAAAGACACAGTAAAGAAATTAAGAGATTCATTT 362
Oy 101 Prolysalagluluglualeuysalagluleuysargtryrthrlysglutrythrly 120
Db 363 CCCAAGCAGAAAGCTGAAGGACAGCTGTTAAACATATACCAAGAAATATACGAA 422
Oy 121 tyraengluyluslylslyglualaglululeualatarganmetalilegln 140
Db 423 TATATGAGAAAGAAAGAAAGAAAGAGAGAGATTTGGCCCGAATGCGCATCCACAA 482
Oy 141 gluileugluylusgluylargvalalagluluglnlunthrhlgluuglu 160
Db 483 GAGCTGAAAAGAAAACAGAGGTGACACAGAGCGAGCATTTGAAACAGAA 542
Oy 161 gluphehlalaphelglumetleargansngluileugluylusgluargleuyls 180
Db 543 CAGTCCATGCTTGTGAGAGATGATCCGAAACAGAGCTAGAAAAGCGCATGAAA 602
Oy 181 llevalglnlupheglylylsvalaspprogluyluglyproleuvalproaspheu 200
Db 603 ATTGACAGAGATTTGGAGAGTGAACCTGCTAGGTGGCCGCTGACTTG 662
Oy 201 GluylsProserleuaspvalpheprothrleuthrvalserseerileglnProserasp 220
Db 663 GAGAGCCCTCTTATGATGCTTCCCACTTAAAGCTCATCATACAGCTTCAAGAC 722
Oy 221 Cyshtlthrvalargproalalyseproprovalalaspargserleuylprogly 240
Db 723 TGTCAACAACCTGTAAGGCCAGCTAAGCCACTGTGTGGACAGGTCTTGAACCTGGA 782
Oy 241 Alaleuseranaserleuserilleprothrilleaspglyleuarghlsvallvalpro 260
Db 783 GCACTGACAACTCAGAAAGATHTTCCCAATCATGATGATGCCCATGTGTGTGCT 842
Oy 261 glyargleucysproglupheleuglualeuylaseralaaanthralaarglyvalglu 280
Db 843 GGGCGGCTGTGGCCACATTTCTCCAGTTACCGAGTCCAACTGCGCGGGGAGTGGAG 902
Oy 281 Thrcysglylleuvcysglylylsleumetargansgluphetrhllethrhlsvallleu 300
Db 903 ACATGTGAATTCCTGTGGAAGAACTGATGAGGAGATTAATTTACATTACCACTTTCTC 962
Oy 301 lleprolysglnseralagliserasptrycysanthrhlgluasnugluuglualeupe 320
Db 963 ATCCCAAGCAAGTGTGGGTCTGATTACTGCAACAGAGAGAGAGAAAGAACTTTTC 1022

Alignment Scores:

QY 321 LeuileglnaspginginglyLeuileThreuglytTpIleHstHrHsiprothrgln 340
Db 1023 CTCATACAGAGATCGAGGGGCTCATACACTGGGCTGAGATTACTTACACCCACAG 1082
QY 341 ThrAlaPheLeuSerSerValAspLeuHstHrHsCysSerTyrglnMetMetLeuPro 360
Db 1083 ACCCGGTTTCTCTCAAGTGTGACCTACACACTACCTCTTCCAGAGAGTGTCCCA 1142
QY 361 GluSerValAlaIleValCysSerProLysPheGlnIuThrglyPhePheLysLeuThr 380
Db 1143 GAGTCAGTACGCACTTGTGTGCTCCCAAGTTCAGGAAACTGGATTCTTTAAACTACT 1202
QY 381 AspIleGlyLeuGlnIuLysSerCysArgGlnIuLysGlyPheHstProHsSerLys 400
Db 1203 GACCATGAGCTAGAGAGATTCTTCTGTGCGCAAGAAAGATTTCATCCACACAGCAAG 1262
QY 401 AspProLeuPheCysSerCysSerHstValThrValAlaAspArgAlaValThrIle 420
Db 1263 GATCACTCTGTCTGTGAGCTGACGCACTGCTGTGTGACAGACAGCAGTGCACATC 1322
QY 421 ThrAspLeuArg 424
Db 1323 ACAGACCTTCCA 1334
RESULT 6
AY083159 1544 bp mRNA linear ROD 26-MAR-2002
LOCUS Rattus norvegicus AMSH mRNA, complete cds.
DEFINITION AY083159
VERSION AY083159.1 GI:19743767
ACCESSION
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1544)
AUTHORS Pawlak, A. and Guellaen, G.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2002) U99, INSERM, Hopital Henri Mondor, Creteil
94010, France
FEATURES
source
location/Qualifiers
1..1544
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
51..1325
/codon_start=1
/product="AMSH"
/protein_id="AA02520.1"
/db_xref="GI:19743768"
/translation="MSDHAIVSLPPQDRVILISQLSAVELNEDIPPRYFRSGVEII
/transl_except="MSDHAIVSLPPQDRVILISQLSAVELNEDIPPRYFRSGVEII
RMSIYSEGNIEHAFILYNYKYLFLIEKLPHRDYSAIIPKXDAVKLKNVAPK
AERIKTLKRYTKEVQYKREKKEEELAPNLAIOELKXKOPVAKXOKOLEE
OFHAFERKTIQKQELBKRLKIVQSFKVDPPGCPPLPDLKRCVDAVSSPSPTCT
SDCNTTLRPAPKPVADSLKPGALSVLENVPTLGLHIVAPNLCSFVQLASANTA
KIGLTCGLCGKLMKNSFTLHVLIPQNGPDPCHENNEELTFMDDDGLTLGNI
HTHPTQAFSLSDLHTHCSYOMLPSIAIVCSPKFQETGFFLIDYGLQISTCPO
KGFPHGSDPFLPCDSHVTWKRIIVITDLR"
ORIGIN
Alignment Scores:
Pred. No.: 5,17e-147 Length: 1544
Score: 1913.00 Matches: 357
Percent Similarity: 91.98% Conservative: 33
Best Local Similarity: 84.20% Mismatches: 34
Query Match: 86.64% Indels: 0
DB: 10 Gaps: 0
US-09-831-452-1 (1-424) x AY083159 (1-1544)
QY 1 MetSerAspHsIGlyAspValSerLeuProGlnuAspArgAlaValAlaLeuSerGln 20
|||||
|||||

Db 51 ANGTCTGACCATGCAAGATGTAGCTCCACCCCAAGACCGGTAGAGATTGTGCGCA 110
QY 21 LeuGlySerAlaValGluValAsnGlnuSerIleProProArgArgTyrrPheArgSerGly 40
Db 111 CTCGGTAGTACGATTGAGTTAATAGACATTCCGCCCGCGCTACTTGTGTCGGGT 170
QY 41 ValGlnIleIleArgMetLaseTletYrSerGlnuGlyAsnIleGlnHsIlePhe 60
Db 171 GTTAGATATATCGATGATGATCATCTTACTGAAAGAGGCAATTGAATGATGCTTT 230
QY 61 IleLeuTyrrAsnLysTyrrIleThrLeuPheIleGlnLysLeuProLysHsIleArgAspTyrr 80
Db 231 ATCTCTTACAAAGATGATCATCGCTGTTATTGAAAAAATTCCAAAACCGAGACTAC 290
QY 81 LysSerAlaValIleProGlnuLysLysAspThrValLysLeuLysGlnIleAlaPhe 100
Db 291 AATCGCCCATCATTCGCCAGAAAGACCGGCTCAAGAAATTAAAGATTGCTTTC 350
QY 101 ProLysAlaGlnuLysLeuLysAlaGlnuLeuLysArgTyrrThrLysGlnuTyrrGlnu 120
Db 351 CTTAAAGCGGAAGAGCTGAAAGACAGACTCTTGAAGAGATACCAAGAGATATGACGAG 410
QY 121 TyrrAsnGlnuLysLysLysGlnuIleGlnuLeuAlaArgAsnMetAlaIleGlnu 140
Db 411 TATAAGAGACGGAAGAGAGAGAGAGAGAACTTGCCGAAATATCGCCATCCAGCA 470
QY 141 GluLeuGlnuLysGlnuLysGlnuArgValAlaGlnuGlnuLysGlnuGlnuGlnu 160
Db 471 GAATCGAAAGAAAG 530
QY 531 CAGTTCAAGCTTTGAGAAAGATGATCCAGAGACAGAGACTGAGAAAGAGCGGCTAAA 590
QY 181 IleValGlnuLysGlnuLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 591 ATTTGTCMAAGATTCGGAGAGTAGACCTGCGCGCGGCGCTGCTGCTGCTGATGTCG 650
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 651 GAAAGCCCTGTGTGATGTGCGCCCACTTACCTTCTCCCAAGAGACTTCAGAG 710
QY 221 CysHstThrThrValArgProAlaLysProProValAlaAspArgSerLeuSerProGly 240
Db 711 TGTAAACAAACCTGAGCGACACTAAAGCACTGTGTGAGACAGGCTCTCAAACTGGA 770
QY 241 AlaLeuSerAsnSerGlnuSerIleProThrIleAspGlyLeuArgHstValValPro 260
Db 771 GCATTAAAGGCTATGAAATGTTCCACCAATTGAAGGCTCGGCCACTTGTGCGCC 830
QY 261 GlyArgLeuCysProGlnuPheLeuGlnuLeuAlaSerAlaAsnThrAlaArgGlyValGln 280
Db 831 CGGAATCTGAGTCAAGAAATTTCTCCAGCTTGCAGAGGCAACACATGCGCAATCGAG 890
QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGlnuPheThrIleHstHsIleValLeu 300
Db 891 ACCTGAGAGTCTCTGTGTGAAAACGAGAGAAATGAATTCACATTCACATGTCTTC 950
QY 301 IleProLysGlnuSerAlaGlySerAspTyrrCysAsnThrGlnuAsnGlnuLysLeuPhe 320
Db 951 ATCCCAAGCAAAATGCTGGGCTGATTTATTCACACAGAGATGAAGAAGAAATTTTC 1010
QY 321 LeuileglnaspginginglyLeuileThreuglytTpIleHstHrHsiprothrgln 340
Db 1011 TTTATCAGAGATGATCTTGATCTCCACTCTGTGTGATCCACACCCATCCAAACCA 1070
QY 341 ThrAlaPheLeuSerSerValAspLeuHstHrHsCysSerTyrglnMetMetLeuPro 360
Db 1071 ACAGCTTTCTGTCCAGTGTGATCTGACACAGCACTGCTCTACCAATGATGTTACCA 1130
QY 361 GluSerValAlaIleValCysSerProLysPheGlnuThrglyPhePheLysLeuThr 380
Db 1131 GAGTCATAGCAATTGTGTGCTCCCAAGTTCAGAGAGACTGAGATTCTTTAAATTAAC 1190

Oy 381 Asph1sglyeugluglul1eserSerCyArGlnlyglyphenisProHisSerIys 400
 Db 1191 GACATATGCGCTTCAAGATATTCAACTGCGCGAGAAAGCTTCAACCCCATGGGAG 1250
 Oy 401 AsprprroleupheCySerSerHisValThrValaAspaGalaValThrIle 420
 Db 1251 GACCACCGCTGTCTGTGACTGCAGCATGTCTGTCAAGACAGAAATTGGAGATC 1310
 Oy 421 ThrAspleuAry 424
 Db 1311 ACAGACCTTGA 1322
 RESULT 7
 BC061711 1735 bp mRNA linear ROD 06-NOV-2003
 LOCUS Rattus norvegicus associated molecule with the SH3 domain of STM,
 DEFINITION mRNA (cDNA clone MGC:72308 IMAGE:5597943), complete cds.
 ACCESSION BC061711
 VERSION BC061711.1 GI:38197539
 KEYWORDS MGC.
 ORGANISM Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1735)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schaller,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,A.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Scieplon,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.U., Usdin,T.B., Teshigahara,S.,
 Canninci,P., Trange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Ackerman,R.D., Mullany,S.C., Bosak,S.A., McEwan,P.J.,
 Moxley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
 Dickinson,N.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
 Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL MEDLINE 22386257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 1735)
 AUTHORS Strausberg,R.
 TITLE Direct Submission
 JOURNAL Submitted (05-NOV-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: gcgaps@email.nih.gov
 Tissue Procurement: Jeff Green/Patruu Kondaiah, NCI.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland.
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Ahlter,N., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,C.L., Mastello,C., Maekeri,B., Mastrian,S.D., McCloskey,J.C.,

FEATURES
 source
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: ISAK Place: 137 Row: a Column: 7
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 19924064.
 Location/Qualifiers
 1..1735
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="MGC:72308 IMAGE:5597943"
 /tissue_type="Prostate, pool of NCI CGAP Pr30, 40, 41 and
 NCI CGAP Pr29, 39, 42"
 /clone_id="NCI CGAP_P149"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6.1"
 1..1735
 /gene="Pamsh"
 /db_xref="locusID:171565"
 /db_xref="RGD:619963"
 227..1501
 /codon_start=1
 /product="Pamsh protein"
 /protein_id="P461711.1"
 /db_xref="GI:38197540"
 /db_xref="locusID:171565"
 /translation="MSDADVSLPPQVRVILISQSAVELNEDIPRRYFRSGVEII
 RMASIVSEGNIEHAFILNYKTLFIEKPKHRYDAIIOGLEKQKQVAKQKQLOBO
 OEHAPEKIKQELKEXRIKQYKRRKEEELARNALIOGLEKQKQVAKQKQLOBO
 SDCTTLAPKPPVYDSLRKGLSVENPTTISGLNHYVPRLCSEPIQLASANTA
 KIGTQGLGCKLKNSEFTYTHVLPKONGPDYCHNESEELTFPMDDJGLTLGWT
 HTHPTQAFYSVLDLTHCSYQMLPESSIAVCSPKQETGFFFLTDYGLQELSTCRQ
 KGFPHGSDPLFDCSHVWVKDRIVITDLR"
 980..1309
 /note="Nov4/MPN/PAD-1 family. Members of
 this family are found in proteasome regulatory subunits,
 eukaryotic initiation factor 3 (eIF3) subunits and
 regulators of transcription factors. This family is also
 known as the MPN domain and PAD-1-like domain. It has been
 shown that this domain occurs in prokaryotes"
 /db_xref="CDP:pfam01398"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,02e-147 Length: 1735
 Score: 1913.00 Matches: 357
 Percent Similarity: 91.98% Conservative: 33
 Best Local Similarity: 84.20% Mismatches: 34
 Query Match: 86.64% Indels: 0
 DB: 10 Gaps: 0
 US-09-831-452-1 (1-424) x BC061711 (1-1735)
 Oy 1 MetSerAsph1sglyaspvalserleupProGluAspaRgValArgalaLeuSergln 20
 Db 227 ATGTGTGACCATGAGAGATGAGCTCCACACCAAGCGGGAGAGATTCGTGCAA 286
 Oy 21 LeuGlySerAlaValGluValAsrGluAsp1leProProArgATgTyPheAsgSerGly 40
 Db 287 CTCGGTAGTCAGTGTGATTAAGAACATTCGCCGCCGCTCTTCGTCGGGT 346
 Oy 41 ValGluIleAlaRgMetAlaSer1lTySerGluGluGluAen1leGluHsAlaPhe 60
 Db 347 GTTGAGATCAACCGCATGCGATTCATTCCTTAAGAAGCAACATTTGAACGCTTT 406
 Oy 61 IleLeuTyAsnlySYrIleThLeuPhe1leGluIysleuPolysHsaigaSPYr 80
 Db 407 ATCTCTACACAGATGATCAGCGCTGTTATTAATAAACAATCCAAACACCGAGACTAC 466

```

QY 81 LysSerAlaValIleProGluLysAspThrValIleLysLeuValIleAlaPhe 100
Db 467 AAATCGGCATCATTTCCCGAAGAAAGACGGCTCAAGAAATTAAAGATGCGCTTTC 526
QY 101 ProLysAlaIleGluLeuLysAlaGluLeuLeuLysArgThrLysGluThrThrGlu 120
Db 527 CCTAAAGCGGAAGAGCTGGAAGACAGAACTCTTAAGAGATACCAAAAGATGAGCGAG 586
QY 121 TyrAsnGluGluLysLysGluValGluGluLeuAlaArgAsnMetAlaIleGln 140
Db 587 TATAAGAGCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 646
QY 141 GluLeuGluLysGluLysGluValAlaGlnGluLysGlnGlnGlnGlnGlnGln 160
Db 647 GAACTGGAAAAAGAAAAAGCAGAGAGTTCACAGCAGAGAGAGAGAGAGAGAGAGAG 706
QY 161 GluPheHisAlaPheGluGluMetIleArgAsnGluGluLysGluValGluLys 180
Db 707 CAGTTCATGCTTGAAGAGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
QY 181 IleValGlnGluPheGluLysValAspProGluLysGluProLysValProAspLeu 200
Db 767 ATTGTTCAAGAGTTCGGGAAGAGTACCTGGCCCGCGGCTGCTGCTGCTGCTGCTG 826
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerIleGlnProSerAsp 220
Db 827 GAAAAGCCGTGTGAGTGTGGCCCGACGTTACCTTTCGCCCGACGAGAGAGAGAGAG 886
QY 221 CysHisThrThrValArgProAlaLysProProValAlaAspArgSerLeuLysProGly 240
Db 887 TGTAAACCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 947 GCATTAGCGCTGAGAAAAATGTTCCACCATTTGAAGAGAGAGAGAGAGAGAGAGAG 1006
QY 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
Db 1007 CGCATCTGCTGCTGAGATTTCTCCAGCTTGCAGGCCCAACAGCTGCCAAGGATCGAG 1066
QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
Db 1067 ACCTGTGAGAGCTCTGTGAGAAACTGATGAGAAATGATTCACAAATCACACATGTTCTC 1126
QY 301 IleProLysGlnSerAlaGlySerAspThrCysAsnThrGluAsnGluGluLeuPhe 320
Db 1127 ATCCCGACAGCAAAATGTGGGCTGATTTATTCACACAGAGATGAAGAAATTTTC 1186
QY 321 LeuIleGlnAspGlnGlnGlnIleThrLeuGluIleThrIleHisThrHisProThrGln 340
Db 1187 TTTATGCAAGATGATCTTGAGCTCTCACTTGTGGCTGATCCACACCACTCAACCA 1246
QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerThrGlnMetMetLeuPro 360
Db 1247 ACCGCCCTTGTGTCAAGTGTGATCTGCACACGACGCTCTTACCAAAAGATGTTACCA 1306
QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGluPhePheLysLeuThr 380
Db 1307 GAGTCCATGAGATTTGCTGCTCCCCCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAG 1366
QY 381 AspiHisGlyLeuGluLysLeuSerSerCysArgGlnLysGluPheHisProHisSerLys 400
Db 1367 GACTATGGCTTCAAGAGATTTCAACTGCGCGAGAGAAAGGCTTTCACCCCGATGGCAGA 1426
QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValIleThrIle 420
Db 1427 GACCCACCGCTGTTGCTGAGCTGAGCGCATCTGTCAAAAGCAAGATTTGACGATC 1486
QY 421 ThrAspLeuArg 424
Db 1487 ACAAGCCTTGA 1498

```

```

RESULT 8
E35542
LOCUS E35542 1384 bp DNA linear PAT 31-JAN-2002
DEFINITION Protein AMSH and cDNA thereof.
ACCESSION E35542
VERSION E35542.1 GI:18624563
KEYWORDS JP 2000139469-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1384)
Sugamura, K. and Tanaka, N.
Protein AMSH and cDNA thereof
Patent: JP 2000139469-A 2 23-MAY-2000;
SCIENCE & TECH AGENCY
COMMENT
OS Homo sapiens (human)
PN JP 2000139469-A/2
PD 23-MAY-2000
PF 12-NOV-1998 JP 1998322674
PR
PT KAZUO SUGAMURA, NOBUYUKI TANAKA
PC C12N15/09, C07K14/52, C07K16/24//C12N1/21, C12P21/02, C12N15/09,
PC C12R1:91,
PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), C12N15/00, (C12N15/00,
PC C12R1:91)
CC
FH Key CDS Location/Qualifiers
FT 56..1327.
FEATURES
source 1..1384
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2,44e-146 Length: 1384
Score: 1904.00 Matches: 354
Percent Similarity: 91.75% Conservative: 35
Best Local Similarity: 83.43% Mismatches: 35
Query Match: 86.23% Gaps: 0
DB: 6
US-09-831-452-1 (1-424) x E35542 (1-1384)
QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
Db 56 ATGTCTGACCATGGGAGATGTGAGCTCCACCCCAAGACGGGTAGAGATTCTGTCCAA 115
QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgThrPheArgSerGly 40
Db 116 CTTGGAGAGCGAGTTGAGTTAATGAAGACATTCACCCCGCGTACTACCGCTCCGGT 175
QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluLysAsnIleGluHisAlaPhe 60
Db 176 GTTGAGATCATCCGATGGCGCTTACTGGAAGAGCAACATTAACATCCCTT 235
QY 61 IleLeuTyrAsnLysThrIleThrLeuPheIleGluLysLeuProLysHisArgAspThr 80
Db 236 ATCTCTTCAACAAGTACATCAAGCTGTTATTTGAAAACTTCGAAACACCGAGAGCTAC 295
QY 81 LysSerAlaValIleProGluLysAspThrValLysLysLeuLysGluIleAlaPhe 100
Db 296 AAATCAGTATCATCTCTGAGAAAGAAAGATGCTGTCAGAAATTAAGAGCGCTGCTTTC 355
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgThrLysGluThrThrGlu 120
Db 356 CCTAAAGCGGAAGAGCTGAGAGACAGAGCTCTTGAGAGATACCAAGAAATTTGAGCAG 415
QY 121 TyrAsnGluGluLysLysGluValGluGluLeuAlaArgAsnMetAlaIleGln 140
Db 416 TATAAGAGCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 475

```


OY		14	GluLeuGlnLysGlnLysGlnArgValAlaGlnGlnLysGlnGlnLysGlnGlnLys	16
Db		476	GAGCTGGAAAAGAAACAGAAGGTTCTCTCAACAAGAACAGACACTTAAGCAGAG	535
OY		161	GlnPheHisAlaPheGlnGlnMetIleArgAsnGlnGlnLeuGlnLysGlnArgLeuLys	180
Db		536	CATTTCACAGCCCTTTCAGAGATGTATCCACAGAGCAGAGACTGGAAAAAAGCGCTGA	595
OY		181	IleValGlnGlnPheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeu	200
Db		596	ATTGTTCAAGAGTTCCGGAGAGTAGACCCTTGCCCCCTGGGGGCTCTGCCTCATCTG	655
OY		201	GlnLysProSerLeuAspValPheProThrIleuThrValSerSerIleGlnProSerAsp	220
Db		656	GAAGAACCTTGTGTGATGTGGCCCCCACACTCAACGCTTCCGCCACGACAGCTCCAGAC	715
OY		221	CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly	240
Db		716	TGTAAACACAGGACATAGCCCACTAAGCCACCTCTGTGTGACAGAGTCCCTGAAACCTGGA	775
OY		241	AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValAlaPro	260
Db		776	GGCTTAAGGCGTATGAAATGTTCCACCATTTGAAGCGCTCCGACACATCTGTGGTCC	835
OY		261	GlyArgLeuCysProGlnPheLeuGlnIleuAlaSerPalaAsnThrAlaArgGlyValGlu	280
Db		836	CGTAATCTGTCTCGAATTTCTTCAGCTTCCAGTCCAGTCCCAATACGCCCAAGACATTTGA	895
OY		281	ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGlnPheThrIleThrHisValLeu	300
Db		896	ACCTGTGGAGTCTCTGTGGAAAACTGATGAGAAATGAATTCACAAATCACATGTTCTC	955
OY		301	IlePolysGlnSerAlaGlySerAspTrpCysAsnThrGluAsnGlnGlnGluLeuPhe	320
Db		956	ATCCCAGACAAATAGTGTGGCCCTGATTATTCACACGGAATAAGAACAAATTTTC	1011
OY		321	LeuIleGlnAspGlnGlnGlyLeuIleThrIleuGlyITripIleHisThrHisProThrgln	340
Db		1016	TTTATTCAGAGATGACTTGGAATCCCACTCTTGGCTGGATCCATACATCAACCCAA	1072
OY		341	ThrAlaPheLeuSerSerValAspIleuHisThrHisCysSerTyrGlnMetSerLeupro	360
Db		1076	ACGGGCTTCTGTCCAAGTGTGATCTCCACACTCACTGCTCTTACCAATGATTTACCA	1133
OY		361	GluSerValAlaIleValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuthr	380
Db		1136	GAGTCATATCGCAATGCTGTCTGCCCAAAGTTCCAGAAACCTGATTCCTTAAGTAACT	1192
OY		381	AspHisGlyLeuGlnGlnLysSerSerCysArgGlnLysGlyPheHisProHisSerLys	400
Db		1196	GACTATGGATCTTCAAGAGATTTCACCTGGCGGAGAAAGGTTTCACCCCCATGGACGA	1255
OY		401	AspProProLeuPheCysSerCysSerHisValThrValValAspArgAlaValThrIle	420
Db		1256	GACCCACCGCTTCTGTGTGATCGACGACCATGTCACTGTCAAGAGACGAATTGTGACGATC	1312
OY		421	ThrAspLeuArg 424	
Db		1316	ACAGACCTTCGA 1327	
<hr/>				
RESULT 9				
BC025111		2094 bp	mRNA	linear
LOCUS				
DEFINITION			Mus musculus Stam binding protein, mRNA (cDNA clone MGC:36212	
VERSION			IMAGE:4194332), complete cds.	
KEYWORDS			BC025111.1 GI:19263555	
SOURCE			MGC.	
ORGANISM			Mus musculus (house mouse)	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;	
			Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.	

REFERENCE	Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Steinem, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, D., Hsieh, F., Diatchenko, L., Martins, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. P., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Udell, T. B., Toshiyuki, S., Garinczi, P., Prange, C., Raha, S. S., Loughellano, N. A., Peters, G. J., Abramson, R. D., Mulhany, S. J., Boeck, S. A., McMan, P. J., McKenna, K. J., Malek, U. A., Gunaratne, P. H., Richard, S., Morley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahy, J., Helton, E., Kettelman, M., Madan, A. C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmitz, J., Myers, R. M., Butcherfield, I. S., Krzywinski, M. I., Skalska, U., Smalins, D. E., Scherch, A., Schein, J. S., Jones, S. O., and Marra, M. A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PMID	12398257
REFERENCE	2 (bases 1 to 2094)
REFERENCE	Strausberg, R.
AUTHORS	Submitted (05-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Louieged, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Naravati, A. N., Gibbs, R. A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Series: IRAX Project: 61 Row: 5 Column: 3 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 17941276. Location/Qualifiers 1..2094 /organism="Mus musculus" /mol_type="mRNA" /strain="FVB/N" /db_xref="taxon:10090" /clone="MGC:36212 IMAGE:419432" /tissue_type="Liver, normal. 5 month old male mouse." /clone_id="NCI CGAP_L19" /lab_host="DH10B" /note="Vector: pCMV-Sport6" 1..2094 /gene="Stambp" /note="synonym: Ameh" /db_xref="LOCUSID:70527" /db_xref="MGI:191777" 51..1325 /codon_start=1 /product="Stam binding protein" /protein_id="AA25111.1" /db_xref="GI:1926356" /db_xref="LOCUSID:70527"
CDS	

/trans1action="MSDPHGDVSPQDRVRLISQSGAVELMEDIPRRYRSGVEII
RMAVSVEBENIENHAPILYKRYITLFLKPKRDYSAIIPKQAVKVLKAVAPK
AEELTELLRRYREYQYERKKEEELAEANAIQOEKKEKQVAAQKQOLQOE
OFAPEPMIOROELKEXRIKIOEFGVDPGCGPILPDLKXCVAVASPSPPOT
PDCGTGRPAKPPYVDSBLKPAQSLVYENPTIEGLRHIVPRLNCSSEFIQASANTA
KGIJTCGLVLCGLMRNEFTITHVILIPQNGSDPYCHTENEELFPMQDDITLGLMI
HTHPTQTLASVLDLTHCSYQWMLPEISIVCSPKFQETGFKLMDYGHQETSTCRQ
KGFPHGRDPPLECDSCSHVTVKDRITVITLRL"

misc_feature
804..1133
/note="Mov34: Region: Mov34/MPN/PAD-1 family. Members of
this family are found in proteasome regulatory subunits,
eukaryotic initiation factor 3 (eIF3) subunits and
regulators of transcription factors. This family is also
known as the MPN domain and PAD-1-like domain. It has been
shown that this domain occurs in prokaryotes"
/db_xref="CDD:pfam01398"

ALIGNMENT SCORES:

Pred. No.: 4,19e-146 Length: 2094
Score: 1904.00 Matches: 354
Percent Similarity: 91.75% Conservative: 35
Best Local Similarity: 83.49% Mismatches: 35
Query Match: 86.23% Indels: 0
DB: 10 Gaps: 0

US-09-831-452-1 (1-424) x BC025111 (1-2094)

QY 1 MeSeraspHisGlyaspValSerLeuProGluaspArgValArgAlaLeuSerGln 20
Db 51 ATGCTACATCATGGGAGATGAGCCTCCACCCCAACCGGGTGAAGATCTGTCCTCA 110
QY 21 LeuGlySerAlaValGluValAsnGluAspLeuProArgArgArgArgSerGly 40
Db 111 CTGGGAGTGCAGTGTAAATGAAGACATTCACCCGCTGCTACTCCCTCGGT 170
QY 41 ValGluLeuLeuArgMetAlaSerLeuTyrSerGluGluValAsnLeuGluHisAlaPhe 60
Db 171 GTTGAATCATCCGACGCGCTGCTTACTCGGAAGAGGACCAATGTAACATGCTTT 220
QY 61 LeuLeuTyrAsnLysTyrLeuThrLeuPheLeuGluLysLeuProLysHisArgAspTyr 80
Db 231 ATCCTTACAAACAAGTACATCACCTGTTATTGAAATACTTCGGAACAACCGAGACTAC 230
QY 81 LysSerAlaValLeuProGluLysLysAspThrValLysLysLeuGluLeuAlaPhe 100
Db 291 AAATCAGCTATCATTTCTGAGAAAGAAAGATGCTGTCAAGAAATTAAGAGCGCTTTC 350
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
Db 351 CCTAAACGGGAAGCTGAGAGACAGAGCTCTTGAGAGATACACCAAGAAATATGAGCAG 410
QY 121 TyrAsnGluGluLysLysGluValGluGluLeuLysAlaArgSerMetAlaLeuGln 140
Db 411 TATTAAGCGAAAGAAAGAAAGAAAGAGAGAACTTCCCGCAATATACCGCATCCGCAA 470
QY 141 GluLeuGluLysGluLysGluLysArgValAlaGluGluLysGluGluGluGluGlu 160
Db 471 GAGTTGAAAAGAAAAGAAAAGAGAGGTTGCTTACAGAAAGAAAGAGCGTGAAGAGAG 530
QY 161 GluPheHisAlaPheGluGluMetLeuArgAsnGluLeuGluLysGluArgLeuLys 180
Db 531 CAATTCATGCTTGAAGAGATGATCCAGAGGACGAGAGCTGGAAGAAAGAGCGCTGAAA 590
QY 181 IleValGluGluLysGluLysValAspProGluLysLeuLysProLeuValProAspLeu 200
Db 591 ATTGTCAAGAGTTGGGAAAGTGAACCTGGCCCTCGCGGCTCTGCTCCCTGACTG 650
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerLeuGlnProSerAsp 220
Db 651 GAAAGAGCTGTGTAGATGTGGCCCGACGACCGTTCGCGCAAGAGAGACTCCAGAC 710
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240

Db 711 TTAACACAGGATGAGGCGACCTAAGCACCTGTGTGAGACAGTCCCTGAAACCTGGA 770
QY 241 AlaLeuSerAspHisGlySerLeuProThrLeuAspGlyLeuArgHisValValAlaPro 260
Db 771 GGGTTAAGGCTATGAAATATTTCCACCATTAAGGCTGCGCACATCGTGGTGCC 830
QY 261 GlyArgLeuCyProGluPheLeuGluLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
Db 831 CGAATCTGTGCTGCAAGATTTCTCCAGCTTGCACATGCAATACCGCAAGAGACTTGA 890
QY 281 ThrCysGlyLeuLeuCyGlyLysLeuMetArgAsnGluPheThrLeuHisValLeu 300
Db 891 ACCTGGGAGTCTCTGTGGGAAACCTGATGAAATGAATTCACATCAACATGTTCTC 950
QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
Db 951 ATCCCAAGCAAAATGCTGGGCTGATTTTCCACAGAGAAATGAAAGAAATTTTC 1010
QY 321 LeuIleGlnAspGluGluGluLysLeuLeuThrLeuGlyTyrPileHisThrHisProThrGln 340
Db 1011 TTTATGACAGATGACCTTGAAGTCTCTGCTGCTGATTCATATCATCAACCCAA 1070
QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGluMetLeuPro 360
Db 1071 ACGGCTTCTGTCCAGTGTGATCTCCACACTCACTGCTCTACCAATGATGTTACCA 1130
QY 361 GluSerValAlaIleValCysSerProLysPheGluGluThrGlyPhePheLeuThr 380
Db 1131 GAGTCATGCAATGATGCTGTTCCCAAAAGTCCAGGAATCGATTTTAACTACTAC 1190
QY 381 AspHisGlyLeuGluGluLysSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
Db 1191 GACTATGCTTCAAGATTTCAACCTGCGCGAAGAAAGCTTACCCCATGCGAGA 1250
QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
Db 1251 GACCCACGCTTCTGTGACATGACAGCATGATGATCAAGAGAAATGTGACGATC 1310
QY 421 ThrAspLeuArg 424
Db 1311 ACAGACTTGA 1322

RESULT 10
AB010123 2104 bp mRNA linear ROD 06-DEC-2001
LOCUS AB010123
DEFINITION Mus musculus mRNA for AMSH, complete cds.
ACCESSION AB010123
VERSION AB010123.1 GI:17385633
KEYWORDS AMSH.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Ishii,N., Owada,Y., Yamada,M., Miura,S., Murata,K., Asao,H.,
Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Ishii,N., Owada,Y., Yamada,M., Miura,S., Murata,K., Asao,H.,
Kondo,H. and Sugamura,K.
Loss of neurons in the hippocampus and cerebral cortex of
AMSH-deficient mice
Mol. Cell. Biol. 21 (24), 8626-8637 (2001)
JOURNAL MEDLINE
PUBMED 11713295
2 (bases 1 to 2104)
AUTHORS Ishii,N., Aizawa,K. and Sugamura,K.
DIRECT SUBMISSION
REFERENCE
JOURNAL Submitted (06-JAN-1998) Kazuo Sugamura, Tohoku University School of
Medicine, Department of Microbiology and Immunology, 2-1
Seiryu-machi, Aoba-ku, Sendai, Miyagi 980-8575, Japan
(E-mail:sugamur@mail.cc.tohoku.ac.jp, Tel:81-22-717-8096,
Fax:81-22-717-8097)
FEATURES
source
1, 2104
/organism="Mus musculus"

gene
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 1..2104
 /gene="mamsh"
 56..1330
 /gene="mamsh"
 /note="associated molecule with the SH3 domain of STRAP"
 CDS

/protein_id="BAB78604.1"
 /cdon_start=1
 /product="AMGH"
 /protein_id="BAB78604.1"
 /db_xref="GI:17385634"
 /translation="MSDHDVSLPPDVRVITSLQSAVLENDIPPRVYGVET
 RMSVYSEBNITNHAFLNYKTYITLFBKPKRDKSAILEKNAVKKSVAPK
 AAEKTELLRYTYEYQYKERKEEBELARNIAIQLEKQVADQKQSLQSE
 QFAPEMIOQRELEKRLKVOEPGVDPGPGPLPLEKQCVDAVSPSPOT
 PDNTRPAPKPPVDRSLKLGALSVENPTIEGRHIVPRLNCESELOASATA
 KGIETGCVLGGKLMNEFTTHVLI PRONGPDYCHTENEBEIFFMODLGLTIGMI
 HTPHOTAFISYDLTHGCSYOMLPESAIAYCSPKFOETGPFKLTGYLQELSTCRQ
 KGRPHRDPPLPDCSHVYKDRITVITDLR"

ORIGIN

Alignment Scores:

Pred. No.: 4,21e-146 Length: 2104
 Score: 1904.00 Matches: 354
 Percent Similarity: 91.75% Conservative: 35
 Best Local Similarity: 83.49% Mismatches: 0
 Query Match: 86.23% Indels: 0
 DB: 10 Gaps: 0

US-09-831-452-1 (1-424) x AB010123 (1-2104)

QY 1 MetSerSpHsGlyAspValSerLeuProGluAspArgValaArgAlaLeuSerGln 20
 DB 56 ATGCTACCAATGGGATGTGAGCTTCCACCCCAAGCCGGTGGAGATTCTGCTCCA 115
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 DB 116 CTGGGAGTGGAGTGGATTAAATGAAGACATTCACCCCGCTACTACCGCTCGGT 175
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluValAsnIleGluHisAlaPhe 60
 DB 176 GTTAGATATCATCCCAAGCGCTCGTTACTCGAAGAAAGCAACATGAACATGACCTT 235
 QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 DB 236 ATCTCTACCAACAGTACATCAGCTGTTATTGAATAAATTCGGAACACCGAGACTAC 295
 QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100
 DB 296 AAATCAGCTATCATCTCTGAGAAAGAAAGATGCTGTCAAGAAATTAAGACGCTGCTTC 355
 QY 101 ProLysAlaGluGluLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
 DB 356 CTTAAAGCGAAGAGCTGAAAGACAGAGCTCTTGAGAAAGATACACCAAGAAATGAGCAG 415
 QY 121 TyrAsnGluGluLysLysGluAlaGluGluLysAlaArgAsnMetAlaIleGlnGln 140
 DB 416 TATAAAGAGCGAAAGAAAGAAAGAGAAAGAACTTCCCGAATATTCGCATTCAGCA 475
 QY 141 GluLeuGluLysGluLysGluArgValAlaGlnGluLysGlnGlnGluGluGlnGlu 160
 DB 476 GAGTTGAGAAAGAAAGAAAGAGGTTGCTCAGCAGAAAGAAAGAGCTTACGAGCGAG 535
 QY 161 GluPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgLys 180
 DB 536 CAATTCATGCTTGAAGAGATATCAAGAGCAGAGAGCTGGAAGAAAGACGCTGAAA 595
 QY 181 IleValGlnGluPheGlyValAspProGlyLeuGlyLysProLeuValProAspLeu 200
 DB 596 ATTGTTCAAGAGTTGCGAAGGTAGACCTCGCCCTCGCGGCTGCTGCTGCTGATCTG 655

QY 201 GluLysProSerLeuAspValPheProThrLeuThrValLeuSerIleGlnProSerAsp 220
 DB 656 GAAAAGCCTGTGTGATGATGAGTGCCGCCAGCTACCGTTCTGCCCCAGCAGACTCCAGAC 715
 QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
 DB 716 TGTAAACAGAGCAGAGGAGCAGCTAAGCACCTGCTGGGACAGAGTCCCTGAAACCTGGA 775

QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValValPro 260
 DB 776 GCGTTAAACGCTCATAGAAATGTTCCACCTTAAGGCGCTGGCCACATCTGGTGCCTC 835

QY 261 GlyArgLeuCysProGlnPheLeuGluIleuAlaSerAlaAsnThrAlaArgGlyValGlu 280
 DB 836 CGTATCTGCTCTCAGATATTTCTCAGGCTTGCCAGTGGCAATACCGCAAGGATGAA 895

QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 DB 896 ACCGTGAGAGCCTCTGGGAAAAGTGAAGAAATGATTAACAATACACACAGTGTCTC 955

QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 DB 956 ATCCCAACAAATAATGCTGGGCTGATTTATGTCACACAGAGATGAGAAATTTTC 1015

QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrPheHisThrHisProThrGln 340
 DB 1016 TTTATGCAAGATGACCTTGGACTCTCACTCTTGGCTGGATCCATCTCAATCCAA 1075

QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360
 DB 1076 ACGGCTTCTGCTCAGTGGATGATCTCACACTACGCTCTCTCAAAAGATTTACCA 1135

QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380
 DB 1136 GAGTCATCGCAATGCTGTTCTCCCAAGTTCCAGAAATGATCTTAACTTA 1195

QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 DB 1196 GACTATGCTCTCAAGAAATTTCAACCTCGCGAAGAAAGCTTTCACCCCATGGCAGA 1255

QY 401 AspProProLeuPheCysSerCysSerHisValThrValValAspArgAlaValThrIle 420
 DB 1256 GACCCACCGCTGTTCTGATGACGACGATCTCACTGCAAGACAGATGTGACGATC 1315

QY 421 ThrAspLeuArg 424
 DB 1316 ACAGACCTTCCA 1327

RESULT 11
 BC003497 2122 bp mRNA linear ROD 03-OCT-2003
 LOCUS Mus musculus Stem binding protein, mRNA (cDNA clone MGC:6953
 DEFINITION IMAGE:3153829), complete cds.
 ACCESSION BC003497
 VERSION BC003497.1 GI:13097542
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Cranista; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2122)
 REFERENCE
 1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.A., Bosak, S.A., McEwen, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.U., Hult, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

Fahay, J., Helton, E., Ketteman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butlerfield, Y. S., Krzywinski, M. J., Skalska, U., Smailus, D. E., Schnerch, A., Schein, J. E., Jones, S. J. and Marra, M. A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2122)
Strausberg, R.
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Division, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guarantee, P. H., Garcia, A. M., Lu, X., Hulyk, S. W., Louised, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Nantavali, A. N., Gibbs, R. A.

FEATURES

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.liml.gov>
Series: IRAC Plate: 5 Row: P Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 17941276.
Location/Qualifiers

1. 2122

/organism="Mus musculus"

/mol_type="mRNA"

/strain="129, C57BL/6J, FVB/N"

/db_xref="taxon:10090"

/clone="MGC:6953 IMAGE:3153829"

/tissue_type="Mammary tumor. Brca1-/E1; MMTV-Cre model. 10 months old. gross tissue."

/clone_lib="NCI CGAP_Mam3"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1. 2122

/gene="Stambp"

/note="synonym: Amsn"

/db_xref="locusid:70527"

/db_xref="MGI:191777"

52. 1326

/codon_start=1

/product="Stam binding protein"

/protein_id="AAH03497.1"

/db_xref="GI:13097543"

/db_xref="locusid:70527"

/translation="MSDHGVSLPPQVRVILISQSAVLNEDIPRRYRGVETIRASVYBEGNIEHAFILVNYKTLFLEKPKADYSALIPKCKDAVKLQVAFPKAEELKTLRLRYTKEYQYERKKKEEELARNALIQELEKQVAVQKQQLDQFAEEMIQOELERKLIVQFEGVDGPGCLPDLKCVDAVAPSPSPQTPQNTGMRPAKPEVDRSLKPGALSVIENVPTTGGALHIVPRLCEPFSQLSANPAKGTCTGATGSKMRNPTTTHVLIPRONGPDCHTENEETFFEMDDIGLITGWIHTPTGVALPSVDLTHSCSYOMLPSIALVCPKQGTGFRITDYGQIBSTGQKGFPHGDPPLFCDSHVTYKDRITVITDLR"

805. 1134

/note="Mov34; Region: Mov34/MN/PAD-1 family. Members of this family are found in proteasome regulatory subunits, eukaryotic initiation factor 3 (eif3) subunits and

misc_feature

ORIGIN

Alignment Scores:

Pred. No.:	4,26e-146	Length:	2122
Score:	1904.00	Matches:	354
Percent Similarity:	91.75%	Conservative:	35
Best Local Similarity:	83.49%	Mismatches:	35
Query Match:	86.23%	Indels:	0
DB:	10	Gaps:	0

US-09-831-452-1 (1-424) x BC003497 (1-2122)

QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
 DB 52 ATGCTGACCATGGGGATGTGAGCTCCACCACCAAGACGGGTGAGATTCTGCTCCAA 111
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 DB 112 CTGGAGAGCGAGTTGAGTTAATGAAGACATTCACCCCGCTACTACCGCTCCGT 171
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGlyAsnIleGluHisAlaPhe 60
 DB 172 GTTGAGATCATCCGATGCGCTCCCTTACTCGGAAGAGCAACATTGAATCCTTT 231
 QY 1IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 DB 232 ACCCTTACAAAGACATCACTACGCTGTTATTGAAAACTTCCGAAACACCGAGCTTC 291
 QY 81 LysSerAlaValIleProGluLysAspThrValLysLeuLysGluIleAlaPhe 100
 DB 292 AATTCAGTATCATCTCTGAGAAAGAGTCCGCTCAAAATTAAAGCCTGCTTTC 351
 QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu 120
 DB 352 CCGAAAGCGGAAGACTGAAGACAGAGCTCTTGAGAGATACACCAAGAAATATGACGAG 411
 QY 121 TyrAsnGluGluLysLysGluAlaGluGluLeuAlaArgAsnMetAlaIleGlnGln 140
 DB 412 TATAAAGACGGAAGAAAGAAAGAGAGGAACTTGCCGAAATATCCCATCAAGCAA 471
 QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnGlnGlnGlnGln 160
 DB 472 GAGTTGAGAAAGAAAGAAAGAGAGGTTGCTCAGCAAGAACAAAGAGCTTACAGCAGAG 531
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgLeuLys 180
 DB 532 CAATTCATGCTTGGAGAGATGATCCAGAGCAGAGAGAGCTGAAAAAGAGCTGAAA 591
 QY 181 IleValGlnGluPheGluLysValAspProGlyLeuGluGlyProLeuValProAspLeu 200
 DB 592 ATTGTTCAAGAGTTCCGAGAGGTAAGCCTGAGCCCTCGGAGCTCTGCTCCTGATGCG 651
 QY 201 GlnLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 DB 652 GAAAGACCTTGATGATGTGCGCCCAAGCTCAGCTTCCGCCACAGCAGACTCCAGAC 711
 QY 221 CysHisThrThrValArgProAlaLysProProValAlaAspSerLeuLysProGly 240
 DB 712 TGTAAACAGGAGATGAGGAGCACTTAAGCCACTGTGTGAGAGAGCTTCTGAACTGCA 771
 QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValValPro 260
 DB 772 GCGTTAAGGCTATGAAATAATGTCCACATTTAAAGCCTGCGCACATCGTGGGCC 831
 QY 261 GlyArgLeuLysProGlnIleLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
 DB 832 CGTAATCTGTGTCAGAAATTTCTCAGGCTTGACAGGCCAAATACGCGCAAGGCAATTGAA 891
 QY 281 ThrCysGlyIleLeuLysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300

regulators of transcription factors. This family is also known as the MN domain and PAD-1-like domain. It has been shown that this domain occurs in prokaryotes"
 /db_xref="CDI:pfam01398"

```

|||||...|||||
Db 892 ACCGTGAGACCTCTGTGAGAAAATGATGAGAAATGAATTCACATCAACATGTTTC 951
Cy 301 ILPEPOLYGLINSEALAGLYSERASPLYCYASATHGULANSNGULUGULPE 320
952 ATCCCAAGACAAATGGTGCCCTGATTATTCACAGAGAAATGAGAAATTTTC 1011
Cy 321 LeuileglnaspnglnglyleuilethleuglyTptlethisthriprothrgln 340
1012 TTTATGACAGATGACCTTGAGCTCTGCTTGCTGATCATCTCATCCACCCAA 1071
Cy 341 ThrAlaheleuserservalaspleuisthrihsysetserytgmletleupro 360
1072 ACCGCCCTTGTGTCATGATGATGATTCACACACCTGCTCCACCAAGATGATCA 1131
Cy 361 GluservAlaAlaIleValCysserProlyspheglnlunthrglyPhePelyLeuThr 380
1132 GAGCCCATCCGCAATCGCTCTGTCCCAAGTCCAGAGAACTGATTTCTTAGCTAATC 1191
Cy 381 AsphlselyleugluglileserSerCyazrglnlysglyPheHsPROHsSeryls 400
1192 GACATATGCTTCAAGATTTCAACCTCGCGAGAAAGCTTTTCACCCCCATGAGCA 1251
Cy 401 AspproProleupheCysserCysserHsValThrValAlaAspArgAlaValThrIle 420
1252 GACCCACCGCTGTTCTGTGACTGACGACCACTGCTACTCTCAAGACAGAAATGTGAGCATC 1311
Cy 421 ThrAspleuArg 424
Db 1312 ACAGACCTCGA 1323
Db 1312 ACAGACCTCGA 1323

RESULT 12
LOCUS BC006939 2172 bp mRNA linear ROD 03-OCT-2003
DEFINITION Mus musculus Stem binding protein, mRNA (cDNA clone MGC:6839
IMAGE:2649718), complete cds.
ACCESSION BC006939
VERSION BC006939.1 GI:13905279
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
MUS musculus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumetazoa; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2172)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, K.H., Spemann, C.M., Schley, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, B., Moore, T., Max, S.T., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Muliahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.D., Malek, J.A., Gunaratne, P.H., Richards, S.,
Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, J., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.T., Skalska, U., Smalins, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, W.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 2172)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgas@dbi.nlm.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisse, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaev, I.,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/INL at: <http://image.llnl.gov>
Series: IRAX Plate: 5 Row: h Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.

FEATURES

source

```

1..2172
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="MGC:6839 IMAGE:2649718"
/tissue_type="Mammary tumor, Brca1-/E1, MMTV-Cre model, 10
months Old, gross tissue."
/clone_id="NCI CGAP_Mam3"
/lab_host="MDH10B"
/notes="vector: pCMV-SPORT6"
1..2172
/gene="Stambp"
/notes="synonym: Ams1"
/db_xref="locusid:70527"
/db_xref="MGI:191777"
107..1381
/codon_start=1
/product="Stambp protein"
/protein_id="AAH06939.1"
/db_xref="GI:13905280"
/db_xref="locusid:70527"
/translation="MSDGDVSLPPQDVRVILSGLSAVELNEDIPRRYRSGVEII
PMASVSEBGNIEHAFILYKVTILFIEKLPKHSDYSALIPKKDAVKKLSVAFPK
AEELKTELRLRYTTEYEQYERKKEEELAPNLAIOBLEKORVAAQOKOLEOE
QFPAFEEIQRQSEKELKIVQFEGVNDPPCPPLDLEKRCVDAVPSPSPTOT
PDCAITGMRPAKPPVADSLKGVASVLENVPTLEGRLIVPRVLCSEFIQLASANTA
KGIETGVLCGKLRNHEFTITHTVILIPQNGPDYCHTNEEEEFPMDDGLTLLGWI
HTHPTGVAFLSVDLHTHGSYOMLPESIAIVCSPKFQETGFFKLTIDYGQIEISTCRO
KGPHHGDPPPLFCDSHVTYKRIYITDLR"
860..1189
/notes="Moy34, Region: Moy34/MYB/PAD-1 family. Members of
this family are found in proteasome regulatory subunits,
eukaryotic initiation factor 3 (eIF3) subunits and
regulators of transcription factors. This family is also
known as the MPN domain and PAD-1-like domain. It has been
shown that this domain occurs in prokaryotes"
/db_xref="CDD:pfam01398"

```

CDS

```

misc_feature
860..1189
/notes="Moy34, Region: Moy34/MYB/PAD-1 family. Members of
this family are found in proteasome regulatory subunits,
eukaryotic initiation factor 3 (eIF3) subunits and
regulators of transcription factors. This family is also
known as the MPN domain and PAD-1-like domain. It has been
shown that this domain occurs in prokaryotes"
/db_xref="CDD:pfam01398"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 4,39e-146 Length: 2172
Score: 1904.00 Matches: 354
Percent Similarity: 91.75% Conservative: 35
Best Local Similarity: 83.49% Mismatches: 35
Query Match: 86.23% Indels: 0
DB: 10 Gaps: 0

```

US-09-831-452-1 (1-424) x BC006939 (1-2172)

Qy	1	MeSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln	20
Db	107	ATGCTGACCATGGAGATGTGAGCCTCCACCCCAAGCCGGGTGAGGATTCCTGCCAA	166
Qy	21	LeuGlySerAlaValGluValAsnGluAspLeuProAlaArgGlyThrPheArgSerGly	40
Db	167	CTTGGAGGTGACGTTGATTAAATGAAGACATTCACCCCGCTACTACCCGCTCGGT	226
Qy	41	ValGluLeuLeuArgMetAlaSerIleTyrSerGluGluGluValIleGluHisAlaPhe	60
Db	227	GTTAGATCATCCGATGCGCTTACTCGAAGAAAGGCAATGTAACATGCGCTTT	286
Qy	61	IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr	80
Db	287	ATCCTTACACAAAGATCATCACCGCTGTTATTGAAAACTCCGAAACACCGAGACTAC	346
Qy	81	LysSerAlaValIleProGluLysLysAspThrValLysLysLeuGluIleAlaPhe	100
Db	347	AAATCAGCTATCATCTCCGAGAAAGAAAGATCTCTCAAGAAATTAAAGAGCGCTTTC	406
Qy	101	ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu	120
Db	407	CCTAAAGCGGAAGAGCTGAAGACAGAGCTTTGAAAGATACCAAGAAATATGACAG	466
Qy	121	TyrAsnGluGluLysLysGluAlaGluGluLeuLysArgAsnMetAlaIleGlnGln	140
Db	467	TATAAGCGCAAGAAAGAAAGAAAGAGAACTTCCGAAATATGCGCATCCAGCAA	526
Qy	141	GluLeuGluLysGluLysGluArgValAlaGlnGlnLysGlnGlnLeuGluGlnGlu	160
Db	527	GAGTTGGAAAAAGAAAACAGAGGCTTCTCAGCAGAAAGCAGAGCTGAGCAGAG	586
Qy	161	GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgLeuLys	180
Db	587	CAATTCCATGCTTTGAGAGATATCATCAGAGCGCAGAGCGTGGAAAAAGACGCTGAAA	646
Qy	181	IleValGlnGluPheGlyLysValAspProGlyLeuGlyProLeuValProAspLeu	200
Db	647	ATTGTTCAAGAGTTCGGAGAGTTCAGCCCTGGCCCTCGGCGCTCTCTCTCTGATCTG	706
Qy	201	GluLysProSerLeuAspValPheProThrLeuThrValSerIleGlnProSerAsp	220
Db	707	GAAAGCCTTGTAGATGTGCCCCCAGCTCACCGTCTCGCCCAAGCAGATCCAGAC	766
Qy	221	CysHisThrThrValArgProAlaLysProProValAlaAspArgSerLeuLysProGly	240
Db	767	TGTACACAGGACATGAGCGCCAGCTTAAGCACTGTGTGTGACAGGTCCCTGAAAACCTGGA	826
Qy	241	AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro	260
Db	827	GCGTTAAGCGTCATAGAAAATGTTCCACCATTTAAGCGCTCGGCCCATCGTGTGCC	886
Qy	261	GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgIleValGlu	280
Db	887	CGTAATCTGTGCTCAGATTCTTCACCTTCCAGTCCCAATACCGCAAGGATGAA	946
Qy	281	ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu	300
Db	947	ACCTGTGAGACCTCTGTGAGAAAATCATAGAAATATTCACAAATCACAATGTTCTC	1006
Qy	301	IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe	320
Db	1007	ATCCCCAGACAAATGTGTGGCGCTGATTATGACACGAGAAATGAAGAAGAAATTTTC	1066
Qy	321	LeuIleGlnAspGlnGlnGluLeuIleThrLeuGlyThrIleHisThrHisProThrGln	340
Db	1067	TTTATGAGGAGATGACTTGAAGCTCTCTCTGTGCTGATCCATCTCATCCAAACCAA	1126
Qy	341	ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro	360
Db	1127	ACGGCCTTTCTGTCTCAGTGTGATCTCCACACTCATGCTCTTAACAAATGATGTTTCCA	1186
Qy	361	GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr	380

Db	1187	GAGTTCATGCATATGCTGTGTTCCCAAGTTCACGAAAGATGATTCCTTAAGTAACT	1246
Qy	381	AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys	400
Db	1247	GACTATGCTTCAACAGATTTCACCTGCGGACGAAAGGCTTTCACCCCATGSCACA	1306
Qy	401	AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle	420
Db	1307	GACCAACCGCTGTCTGTCTACTGACAGCATGTCATGTCACAGACAGATTTGACGATC	1366
Qy	421	ThrAspLeuArg 424	
Db	1367	ACAGACCTTGA 1378	

REFERENCE

AUTHORS

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

RESULT 13

BC055512

Danto rerio cDNA clone MGC:16147 IMAGE:5410887, complete cds.

BC055512

BC055512.1 GI:33416606

MGC.

Danto rerio (zebrafish)

Danto rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danto.

1 (bases 1 to 2455)

Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carinci, P., Plange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Guneratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,

Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, I.S., Krzyzanski, M.I., Skalska, U., Small, D.B.,

Schnerch, A., Schein, J.E., Jones, S.J., and Matra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

2388257

12477932

2 (bases 1 to 2455)

Straussberg, R.

Direct Submission

Submitted (01-AUG-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIN-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Sumio Sugano

cDNA Library Preparation: Dr. Sumio Sugano

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.sngc.stanford.edu>Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LINT at: <http://image.llnl.gov>

REFERENCE

1. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
- Kikuchi, K., Ishii, N., Asao, H. and Sugamura, K.
Identification of AMSH-LP containing a Jab1/MPN domain metalloenzyme motif
Biochem. Biophys. Res. Commun. 306 (3), 637-643 (2003)
- JOURNAL
MEDLINE
PUBMED
12810066
2 (bases 1 to 1601)
Ishii, N., Kuroda, H. and Sugamura, K.
Direct Submission
Submitted (06-JAN-1998) Kazuo Sugamura, Tohoku University School of Medicine, Department of Microbiology and Immunology; 2-1 Seiryō-machi, Aoba-ku, Sendai, Miyagi 980-8575, Japan (E-mail: sugamura@mail.cc.tohoku.ac.jp, Tel:81-22-717-8096, Fax:81-22-717-8097)

FEATURES

SOURCE

CDS

Location/Qualifiers
1..1601
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue="peripheral blood lymphocytes"
34..1344
/codon_start=1
/product="AMSH-LP"
/protein_id="BAC77766.1"
/db_xref="GI:31960350"
/translation="MDQPTVYENGLKLAAMPDHTVLSPEERVALSKGCTITISE
DIPRRYFSGVEMERMAVYLEGNLENKAVLYNKRFTLVKLEPNHSDYQCAVPE
KODIMKLEIAFPRIIDELKNDLKRYNVOYLOSKNYKAEILKKEHOLIEAE
KRRIAKRQOOLESEDFEPEDLKQELARQMSQOTGLSEQIDGSLSCFSTHQ
NNSLIVFADQPKNSDATNVAASHSPVNRALTPAATSAQONLVVGLGCVLPEDLC
HKRIQLAESTVTRGIERFCGLCKGLTNEFTIHTVTPROSAQPDQMDENVELEFNV
ODQDLTLGWIHTPIQTAPLSVDHTICSTOLMPEKIAIVCSPKHKDITRIFLT
NAGLEVSACKKGFHPHTEPRFISICKVILVDIKITLVDLR"
1601
/note="20 A nucleotides"

ORIGIN

Alignment Scores:

Pred. No.: 2,35e-94 Length: 1601
Score: 1269.50 Matches: 240
Percent Similarity: 73.02% Conservative: 74
Best Local Similarity: 55.81% Mismatches: 101
Query Match: 57.50% Indels: 15
Gaps: 2

US-09-831-452-1 (1-424) x AB010120 (1-1601)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValAlaGalaLeuSerGln 20
DB 79 ATGCCTGAGCCATACAGATGTTCCCTTAAGCCAGAGCGAGTCCGCTTAAGCAAG 138
QY 21 LeuGlySerAlaValGluValAsnGluAspLleProProArgArgTyrPheArgSerGly 40
DB 139 CTTGCTGTAAATATACACCATCAGAGAACATCACTCCACGACCTTAAGGTCCTGA 198
QY 41 ValGluIleIleArgMetAlaSerLleTyrSerGluGluGluAsnIleGluHisAlaPhe 60
DB 199 GTAAGAGTGCAGAGATGCGCTCTGTGATTGGAAGAAGAAATTGGAAAATGCTTT 258
QY 61 IleLeuTyrAsnLysTyrLleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
DB 259 GTTCTTATATAATTAATTATACCTTATTTGTAGAAAAGCTTCTTAACATCAGAGTTAC 318
QY 81 LysSerAlaValLleProGluLysLeuAspThrValLysLysLeuLysGluIleAlaPhe 100
DB 319 CAGCAATGTGCAGTACTGAAAGACAGATTTATGAGAAATCGAAAGGATTCATTC 378
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGlu 120
DB 379 CCAAGGACAGATGATTAATAAAGACACTTTTAAAGAAATATACGTAGATACCAAGAA 438

QY 121 TyrAsnGluGluLysLysGluLysAlaGluGluLeuAlaArgAsnMetAlaIleGlnGln 140
DB 439 TATTTCGAAACCAAAACCAATATATAAGCTAAATTTCTCAAAAAATTGACATCAGCA 498
QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGluLysGlnGlnGlnGlnGlu 160
DB 499 TTGATAGAGGACGAAAGAGACGATTCCTAGATTCGCCGACGACGATGAAATCGAG 558
QY 161 GlnPheHisAlaPheGluMetLleArgAsnGlnGluLysGluLysArgLeuLys 180
DB 559 CAGTTTCGTTTTCGAAAGATCAACTCAGAGAGGAGGTTTAGGCCAGGTCGAATCGCA 618
QY 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
DB 619 AGTCAGCAACCTCAGG-----CTCTGACAGACATTC 651
QY 201 GluLysProSerLeuAspValPheProThr-----LeuThrValSer 214
DB 652 GATGGAGCGCTTGTCTGCTTTTCACACACCAACCAATTCCTGCTGAATGATATT 711
QY 215 SerLleGlnProSerAspCysHisThrThrValArgProAlaLysProProValValAsp 234
DB 712 GCAGATCAACCTTAATAAAGTGAAGCAACCAATTATGCTACGCACTCTCTCTGTAAC 771
QY 235 ArgSerLeuLysProGlyAlaLeuSerAsnSerGlnSerLleProThrIleAspGlyLeu 254
DB 772 AGGCGCTTAAGCCACGCTGACTACTAGTCTGTGAGATTAGTGTTGAAGGACATG 831
QY 255 ArgHisValValValProGlyArgLeuGlyProGlnPheLeuGlnLeuAlaSerAlaAsn 274
DB 832 CATGTGTAGTTTGTCCGAAAGATCTTTGCCAATAATTCGCAACTGCGCAATCTAT 891
QY 275 ThrAlaArgGlyValGlnThrCysGlyLleLeuGlyLysLeuMetLysArgAsnLysPhe 294
DB 892 ACGAGGAGAGATGAAACCTGTGGAATCTCTGTGSAATAACGACATTAATGAATTT 951
QY 295 ThrIleThrHisValLeuIleProLysGlnSerAlaGlySerAspTyrCysAsnThrGlu 314
DB 952 ACTATTACCACTGATGATGTGCCAAAGCAGCTGTGGGACCAACCTATTGTGACATGAG 1011
QY 315 AsnGluGluGluLeuPheLeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrPhe 334
DB 1012 AATGAGAGGAATTAATTAATGATGAGATCAACATGATCTCTCCTACCAACATTAAGACACT 1071
QY 335 HisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeuHisThrHisCysSer 354
DB 1072 CATACACATCCACCAACATGCAATTTTATCCACGCTGTGATCTTACACATCAGCTGCC 1131
QY 355 TyrGlnMetMetLeuProGluSerValAlaIleValCysSerProLysPheGlnGlnThr 374
DB 1132 TATCAACTCATGTGCTCCAGAGGCACTTGCATGTTGCTCTACCAACATTAAGACACT 1191
QY 375 GlyPhePheLysLeuThrAspHisGlyLeuGluGluIleSerSerCysArgGlnLysGly 394
DB 1192 GGCATCTTCAGGCTCAACATGCTGCAATGCTGAGTTCGCTGTGAAAAAAGGCG 1251
QY 395 PheHisProHisSerLysAspProLeuPheCysSerCysSerHisValThrValVal 414
DB 1252 TTTTCATCCACACCAACAGAGCCGAGGCTGTTCAGTATATGCCAAACATGTGTGTTAAA 1311
QY 415 AspArgAlaValThrLleThrAspLeuArg 424
DB 1312 GACATAAAAAATATGTGTGATCTGAGG 1341

Search completed: August 23, 2004, 20:50:32
Job time : 5392 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 23, 2004, 19:11:42 ; Search time 4330 Seconds

(without alignments)
3445.477 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208

Sequence: 1 MSBHDVSLPDRVRRLSQ.....FCSCSHYTVYDRAVITIDLR 424

Scoring table:

Xgapop 10.0 ,	Xgapext 0.5
Xgapop 10.0 ,	Xgapext 0.5
Fgapop 5.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 3757730 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xhl
-O=/cgn2_1/USPFO.spool/US09831452/pnauc_17082004_151201_23701.spp query.fasta.1.583
-DB=Pending_Patents_NA_Main -QFMT=fastcap -SUFFIX=p2n.rnps -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HBASESIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09831452 @CGN 1.1 4227 @runat_17082004_151201_23701
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSEPILOCK=100
-LONGLOG -DPV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents_NA_Main:*

1:	/cgn2_6/ptodata/2/pna/US0966_COMB.seq:*
2:	/cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3:	/cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4:	/cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5:	/cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6:	/cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7:	/cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8:	/cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9:	/cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10:	/cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11:	/cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12:	/cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13:	/cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14:	/cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15:	/cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16:	/cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17:	/cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18:	/cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19:	/cgn2_6/ptodata/2/pna/US095_COMB.seq:*
20:	/cgn2_6/ptodata/2/pna/US096_COMB.seq:*
21:	/cgn2_6/ptodata/2/pna/US097_COMB.seq:*
22:	/cgn2_6/ptodata/2/pna/US098_COMB.seq:*
23:	/cgn2_6/ptodata/2/pna/US099_COMB.seq:*
24:	/cgn2_6/ptodata/2/pna/US100_COMB.seq:*
25:	/cgn2_6/ptodata/2/pna/US101_COMB.seq:*
26:	/cgn2_6/ptodata/2/pna/US102_COMB.seq:*
27:	/cgn2_6/ptodata/2/pna/US103_COMB.seq:*
28:	/cgn2_6/ptodata/2/pna/US104_COMB.seq:*

29:	/cgn2_6/ptodata/2/pna/US0966_COMB.seq:*
30:	/cgn2_6/ptodata/2/pna/US0977_COMB.seq:*
31:	/cgn2_6/ptodata/2/pna/US0978_COMB.seq:*
32:	/cgn2_6/ptodata/2/pna/US0979_COMB.seq:*
33:	/cgn2_6/ptodata/2/pna/US0980_COMB.seq:*
34:	/cgn2_6/ptodata/2/pna/US0981_COMB.seq:*
35:	/cgn2_6/ptodata/2/pna/US0982_COMB.seq:*
36:	/cgn2_6/ptodata/2/pna/US0983_COMB.seq:*
37:	/cgn2_6/ptodata/2/pna/US0984_COMB.seq:*
38:	/cgn2_6/ptodata/2/pna/US0985_COMB.seq:*
39:	/cgn2_6/ptodata/2/pna/US0986_COMB.seq:*
40:	/cgn2_6/ptodata/2/pna/US0987_COMB.seq:*
41:	/cgn2_6/ptodata/2/pna/US0988_COMB.seq:*
42:	/cgn2_6/ptodata/2/pna/US0989_COMB.seq:*
43:	/cgn2_6/ptodata/2/pna/US1004_COMB.seq:*
44:	/cgn2_6/ptodata/2/pna/US1005_COMB.seq:*
45:	/cgn2_6/ptodata/2/pna/US1011_COMB.seq:*
46:	/cgn2_6/ptodata/2/pna/US1012_COMB.seq:*
47:	/cgn2_6/ptodata/2/pna/US1013_COMB.seq:*
48:	/cgn2_6/ptodata/2/pna/US1014_COMB.seq:*
49:	/cgn2_6/ptodata/2/pna/US1015_COMB.seq:*
50:	/cgn2_6/ptodata/2/pna/US1016_COMB.seq:*
51:	/cgn2_6/ptodata/2/pna/US1017_COMB.seq:*
52:	/cgn2_6/ptodata/2/pna/US1018_COMB.seq:*
53:	/cgn2_6/ptodata/2/pna/US1019_COMB.seq:*
54:	/cgn2_6/ptodata/2/pna/US1020_COMB.seq:*
55:	/cgn2_6/ptodata/2/pna/US1021_COMB.seq:*
56:	/cgn2_6/ptodata/2/pna/US1022_COMB.seq:*
57:	/cgn2_6/ptodata/2/pna/US1023_COMB.seq:*
58:	/cgn2_6/ptodata/2/pna/US1024_COMB.seq:*
59:	/cgn2_6/ptodata/2/pna/US1025_COMB.seq:*
60:	/cgn2_6/ptodata/2/pna/US1026_COMB.seq:*
61:	/cgn2_6/ptodata/2/pna/US1027_COMB.seq:*
62:	/cgn2_6/ptodata/2/pna/US1028_COMB.seq:*
63:	/cgn2_6/ptodata/2/pna/US1029_COMB.seq:*
64:	/cgn2_6/ptodata/2/pna/US1030_COMB.seq:*
65:	/cgn2_6/ptodata/2/pna/US1031_COMB.seq:*
66:	/cgn2_6/ptodata/2/pna/US1032_COMB.seq:*
67:	/cgn2_6/ptodata/2/pna/US1033_COMB.seq:*
68:	/cgn2_6/ptodata/2/pna/US1034_COMB.seq:*
69:	/cgn2_6/ptodata/2/pna/US1035_COMB.seq:*
70:	/cgn2_6/ptodata/2/pna/US1036_COMB.seq:*
71:	/cgn2_6/ptodata/2/pna/US1037_COMB.seq:*
72:	/cgn2_6/ptodata/2/pna/US1038_COMB.seq:*
73:	/cgn2_6/ptodata/2/pna/US1039_COMB.seq:*
74:	/cgn2_6/ptodata/2/pna/US1040_COMB.seq:*
75:	/cgn2_6/ptodata/2/pna/US1041_COMB.seq:*
76:	/cgn2_6/ptodata/2/pna/US1042_COMB.seq:*
77:	/cgn2_6/ptodata/2/pna/US1043_COMB.seq:*
78:	/cgn2_6/ptodata/2/pna/US1044_COMB.seq:*
79:	/cgn2_6/ptodata/2/pna/US1045_COMB.seq:*
80:	/cgn2_6/ptodata/2/pna/US1046_COMB.seq:*
81:	/cgn2_6/ptodata/2/pna/US1047_COMB.seq:*
82:	/cgn2_6/ptodata/2/pna/US1048_COMB.seq:*
83:	/cgn2_6/ptodata/2/pna/US1049_COMB.seq:*
84:	/cgn2_6/ptodata/2/pna/US1050_COMB.seq:*
85:	/cgn2_6/ptodata/2/pna/US1051_COMB.seq:*
86:	/cgn2_6/ptodata/2/pna/US1052_COMB.seq:*
87:	/cgn2_6/ptodata/2/pna/US1053_COMB.seq:*
88:	/cgn2_6/ptodata/2/pna/US1054_COMB.seq:*
89:	/cgn2_6/ptodata/2/pna/US1055_COMB.seq:*
90:	/cgn2_6/ptodata/2/pna/US1056_COMB.seq:*
91:	/cgn2_6/ptodata/2/pna/US1057_COMB.seq:*
92:	/cgn2_6/ptodata/2/pna/US1058_COMB.seq:*
93:	/cgn2_6/ptodata/2/pna/US1059_COMB.seq:*
94:	/cgn2_6/ptodata/2/pna/US1060_COMB.seq:*
95:	/cgn2_6/ptodata/2/pna/US1061_COMB.seq:*
96:	/cgn2_6/ptodata/2/pna/US1062_COMB.seq:*
97:	/cgn2_6/ptodata/2/pna/US1063_COMB.seq:*
98:	/cgn2_6/ptodata/2/pna/US1064_COMB.seq:*
99:	/cgn2_6/ptodata/2/pna/US1065_COMB.seq:*
100:	/cgn2_6/ptodata/2/pna/US1066_COMB.seq:*
101:	/cgn2_6/ptodata/2/pna/US1067_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	1591	20	Sequence 579, App
2	2208	100.0	1591	20	Sequence 7586, Ap
3	2208	100.0	1767	46	Sequence 32565, A
4	2208	100.0	1910	34	Sequence 2, Appl
5	2208	100.0	1910	53	Sequence 2, Appl
6	2208	100.0	2001	32	Sequence 8614, Ap
7	2208	100.0	2003	55	Sequence 1, Appl
8	2208	100.0	2045	36	Sequence 3931, Ap
9	2208	100.0	2045	43	Sequence 3931, Ap
10	2208	100.0	2087	77	Sequence 4025, Ap
11	2208	100.0	2087	84	Sequence 6951, Ap
12	2208	100.0	2087	89	Sequence 16983, A
13	2208	100.0	2091	46	Sequence 33766, A
14	2208	100.0	2541	25	Sequence 4875, Ap
15	2208	100.0	2541	29	Sequence 3220, Ap
16	2208	100.0	2541	29	Sequence 10797, A
17	2208	100.0	2541	31	Sequence 2597, Ap
18	2208	99.8	1767	102	Sequence 11696, A
19	2208	99.8	1767	102	Sequence 7324, Ap
20	2208	99.8	1767	102	Sequence 7324, Ap
21	2208	99.8	1767	102	Sequence 3912, Ap
22	2208	99.8	1767	103	Sequence 3912, Ap
23	2208	99.8	1767	103	Sequence 7324, Ap
24	2208	99.8	2091	102	Sequence 11695, A
25	2208	99.8	2091	102	Sequence 7323, Ap
26	2208	99.8	2091	102	Sequence 7323, Ap
27	2208	99.8	2091	103	Sequence 3911, Ap
28	2208	99.8	2091	103	Sequence 3911, Ap
29	2208	99.8	2091	103	Sequence 7323, Ap
30	2208	99.8	2091	103	Sequence 7323, Ap
31	2182	98.8	2075	72	Sequence 5172, Ap
32	2182	98.8	2075	72	Sequence 5172, Ap
33	2164	98.0	1850	46	Sequence 32565, A
34	2164	98.0	1850	46	Sequence 11697, A
35	2159	97.8	1850	102	Sequence 7325, Ap
36	2159	97.8	1850	102	Sequence 7325, Ap
37	2159	97.8	1850	102	Sequence 3913, Ap
38	2159	97.8	1850	103	Sequence 3913, Ap
39	2159	97.8	1850	103	Sequence 7325, Ap
40	1913	86.6	1544	1	Sequence 2547, Ap
41	1913	86.6	1544	49	Sequence 2547, Ap
42	1904	86.2	1384	34	Sequence 4, Appl
43	1904	86.2	1384	53	Sequence 4, Appl
44	1882.5	85.3	1797	45	Sequence 21654, A
45	1882.5	85.3	1797	93	Sequence 21654, A

ALIGNMENTS

RESULT 1
US-09-471-275-579
Sequence 579, Application US/09471275
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: Novel Contigs Obtained
TITLE OF INVENTION: From Various Libraries
FILE REFERENCE: 782
CURRENT APPLICATION NUMBER: US/09/471,275
EARLIER APPLICATION NUMBER: US/09/235,076
EARLIER APPLICATION NUMBER: US/09/234,611
EARLIER APPLICATION NUMBER: US/09/240,371
EARLIER APPLICATION NUMBER: US/09/277,227
EARLIER APPLICATION NUMBER: US/09/271,490
EARLIER APPLICATION NUMBER: US/09/253,972
EARLIER APPLICATION NUMBER: US/09/274,861
EARLIER FILING DATE: 1999-03-23
EARLIER APPLICATION NUMBER: US/60/125,453
EARLIER FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: US/60/126,605
EARLIER FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: US/09/306,350
EARLIER FILING DATE: 1999-05-07
EARLIER APPLICATION NUMBER: US/09/399,720
EARLIER FILING DATE: 1999-09-21
EARLIER APPLICATION NUMBER: US/09/404,284
EARLIER FILING DATE: 1999-09-21
EARLIER APPLICATION NUMBER:
EARLIER FILING DATE: 1999-12-16
NUMBER OF SEQ ID NOS: 10451
SOFTWARE: pc_ct_genes Version 1.0
SEQ ID NO 579
LENGTH: 1591
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (199)...(1472)
OTHER INFORMATION: similar to gi4098124 in the genepept database release 114,
US-09-471-275-579
Alignment Scores:
Pred. No.: 1.49e-173 Length: 1591
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-831-452-1 (1-424) x US-09-471-275-579 (1-1591)
1 MetSerAspHisGlyAspValSerIeuProGluAspArgValArgAlaLeuSerGln 20
199 AGGTCTGACATGAGATGATGAGCTCCGCCGGAAGCCGAGGAGGCTCTCCAG 258
21 LeuGlySerAlaValGluValAsnGluAspIleProPheArgGlyPheArgSerGly 40
259 CTGGGAGTGGCGGTGAGGTGATGATGACATTCACCCCTCGTACTTCGCTTGA 318
41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluAsnIleGluAlaPhe 60
319 GTTGAGATTATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 378
61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluIleuProIleuValArgAspTyr 80
379 ACCCTCTATACAAATGATATACCGCTCTTATGAGAACTACCAAAACATGAGATTAC 438
81 IysSerAlaValIleProGluIleuIleuValIleuValIleuValIleuValIleuVal 100
439 AAATCTGCTGTCATTCCTGAAAGAAAGAACACAGTAAAGAAATTAAAGGATTCATTT 498

Qy	101	ProLysAlaGluGluLeuLeuValaGluLeuLeuLysArgTyrThrLysGluTyrThrGln	120
Db	499	CCAAAGCAGAGAGACCTGAGGCGAGACTGTTAAACATATACCAAGAAATATACAGA	558
Qy	121	TyrAenGluGluLysLysGluLualaGluGluLeuLualaArgAspMetValaIleGln	140
Db	559	TATATACAGAAAGAAAGAGAACACAGAGATATGGCCCGGAATCATGCCATCCAGAA	618
Qy	141	GluLeuGluLysGluLysGlnArgValaIaGlnGlnLysGlnGlnGluGlnGln	160
Db	619	GAGCTGGAAAAAGAAAAACAGAGGGTTACACACAGAGACACCAATTGGAAACGGA	678
Qy	161	GlnPheHisAlaPheGluGluMetIleArgAenGlnGluLeuGluLysGluLys	180
Db	679	CAGTTCATGCTTGGAGGAGTATGATCCGGAACACAGAGCTTGAAGAAAAAGCGACTGAA	738
Qy	181	IleValGlnGluPheGluLysValaLaspProGluLysGluLysProLeuValProAspLeu	200
Db	739	ATTGACAGGAGATTGGAGAGTAAACCTGGCCAGAGTGGCCCTGATGCTCACTG	798
Qy	201	GluLysProSerLeuAspValaPheProThrLeuThrValaLysSerIleGlnProSerAsp	220
Db	799	GAGAGACCTCTCTTGAATGATGTTCCCCACCTTAAAGTCTCATCATACAGCTTCAAG	858
Qy	221	CysHisThrThrValaArgProAlaLysProValaLaspArgSerLeuLysProGly	240
Db	859	TGTCAACACACTGTAAAGCCACACTAAGCCACTGTGGTGGACAGCTCTTGAAACTCGA	918
Qy	241	AlaLeuSerAsnSerGlnSerIleLeuProThrIleAspGluLysArgHisValaValaPro	260
Db	919	GCACGACCACTCGAAGATGTTCCCAATCGATGGATGGCGCAATGGTGGTGGCT	978
Qy	261	GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrValaArgGlyValaGlu	280
Db	979	GGGGGGCTGTGGCCACAGTTTCTCCAGTTAGCCAGTGGCAACACTGCCGGGGAGTGGAG	1038
Qy	281	ThrCysGlyIleLeuCysGlyLysLeuMetArgaGlnPheThrIleThrHisValaLeu	300
Db	1039	ACATGTGAATTCCTGTGGAAAACTGATGAGAGATGAATTTAACCATTAACCATTTCTC	1098
Qy	301	IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe	320
Db	1099	ATCCCAAGCAAAAGTCTGGGCTGATTAATCTCAACACAGAAACGAAAGAACTTTTC	1158
Qy	321	LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln	340
Db	1159	CTCATACAGAAATCACAGAGGCTCTATCACACTGGGCTGATTCATATCACACCCCAACAG	1218
Qy	341	ThrAlaPheLeuSerSerValaAspLeuHisThrHisCysSerTyrGluMetLeuPro	360
Db	1219	ACCGGCTTCTTCCAGTGTGCACTACACACTCACTGCTTTCACAGATGATGTGGCA	1278
Qy	361	GluSerValaIaIleValaCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr	380
Db	1279	GAGTCAGTAGCATTGTGTGCTCCCCCAAGTTCAGGAAGTTCGATTCTTTAACTAACT	1338
Qy	381	AspHisGlyLeuGluGluLysSerSerCysArgGlnLysGlyPheHisProHisSerLys	400
Db	1339	GACCACTGACCTAGAGAGATTCTTCCTGTCGCCAGAAAGATTATCAACACAGCAAG	1398
Qy	401	AspProProLeuPheCysSerCysSerHisValaThrValaLaspArgAlaValaThrIle	420
Db	1399	GATCAACCTCTGTTCTGTAGCTGCAAGCCACGTAAGTGTGTGGACAGAGCACTAGCATC	1458
Qy	421	ThrAspLeuArg 424	
Db	1459	ACAGACCTTGA 1470	

```

1  APPLICANT: Yuanhua T. Tang
2  APPLICANT: John Tilligbaast
3  APPLICANT: Ankura Sanku
4  APPLICANT: Chenghua Liu
5  APPLICANT: Radoje T. Dimanac
6  TITLE OF INVENTION: Novel Contigs Obtained
7  TITLE OF INVENTION: From Various Libraries
8  FILE REFERENCE: 784
9  CURRENT APPLICATION NUMBER: US/09/488,725B
10 CURRENT FILING DATE: 2000-01-21
11 PRIOR APPLICATION NUMBER: US 09/004,182
12 PRIOR FILING DATE: 1998-01-07
13 PRIOR APPLICATION NUMBER: US 09/034,341
14 PRIOR FILING DATE: 1998-02-13
15 PRIOR APPLICATION NUMBER: US 09/045,400
16 PRIOR FILING DATE: 1998-03-20
17 PRIOR APPLICATION NUMBER: US 09/321,214
18 PRIOR FILING DATE: 1999-05-26
19 PRIOR APPLICATION NUMBER: US 09/131,598
20 PRIOR FILING DATE: 1998-08-10
21 PRIOR APPLICATION NUMBER: US 09/170,294
22 PRIOR FILING DATE: 1998-10-13
23 PRIOR APPLICATION NUMBER: US 09/179,473
24 PRIOR FILING DATE: 1998-10-27
25 PRIOR APPLICATION NUMBER: US 09/181,430
26 PRIOR FILING DATE: 1998-10-28
27 PRIOR APPLICATION NUMBER: US 09/235,076
28 PRIOR FILING DATE: 1999-01-20
29 PRIOR APPLICATION NUMBER: US 09/234,611
30 PRIOR FILING DATE: 1999-01-22
31 Remaining SEQ Application data removed - See File Wrapper or PALM.
32 NUMBER OF SEQ ID NOS: 10289
33 SOFTWARE: pc_ct_genes Version 1.01
34 SEQ ID NO 7586
35 LENGTH: 1591
36 TYPE: DNA
37 ORGANISM: Homo sapiens
38 FEATURE:
39 NAME/KEY: misc_feature
40 LOCATION: (199)..(1472)
41 OTHER INFORMATION: similar to g14098124 in the genepept database release 114
42 OTHER INFORMATION: Run with FASTX 3.3f00, default parameters
43 US-09-488-725B-7586
44
45 Alignment Scores:
46 Pred. No.: 1,496-173 Length: 1591
47 Score: 2208.00 Matches: 424
48 Percent Similarity: 100.00% Conservative: 0
49 Best Local Similarity: 100.00% Mismatches: 0
50 Query Match: 100.00% Indels: 0
51 DB: 20 Gaps: 0
52
53 US-09-831-452-1 (1-424) x US-09-488-725B-7586 (1-1591)
54
55 QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
56 DB 199 ATGCTCAACCATGAGATGTGAGCTCCGCCCGAAGACCGGGTAGGGCTCTCCACG 258
57 QY 21 LeuGlySerAlaValGluValaGluGluAspGluProProGluArgTyrPheArgSerGly 40
58 DB 259 CTGGGTGTCGGGTAGGAGGTGAATGAACATTCACCCCGGTGTTCCGCTCGGA 318
59 QY 41 ValGluLeuLeuArgMetAlaSerLeuTyrSerGluGluGluValAsnIleGluHisAlaPhe 60
60 DB 319 GTTAGATTATCCGATGGCATCCATTACTCTGAGGAAGCAATGAAACATGCCCTTC 378
61 QY 61 IleuTyrAsnLysTyrIleThrIeuPheIleGluLysLeuProLysHisArgAspTyr 80
62 DB 379 ATCTCTATTAACAAGATATATCAAGCTCTTATTATGGAACATCCAAAACATCGAGATTAC 438
63 QY 81 LysSerAlaValAlleProGluLysAspThrValLysLysLeuLysGluIleAlaPhe 100
64 DB 439 AAATCTCTGCTCTCTCTGAAAAGAAAGACAGTAAAGAAATTAAGGAGATTGCATT 498

```

QY 101 ProTysalaglululeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
 Db 499 CCCAAGACAGAGAGCTGAGAGCGACAGCTGTTAAAGATATACCAAGATATACAGAA 558
 QY 121 TyrAsnGluGluLysLysLysGluAlaGluGluLeuLeuLysAsnMetAlaIleGlnGln 140
 Db 559 TATATGAG 618
 QY 141 GluLeuGluLysGluLysGluLysArgValAlaGlnGlnLysGlnGlnGlnGlnGln 160
 Db 619 GAGCTGAG 678
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluLysGluLys 180
 Db 679 CAGTTCATGCTTCGAGAGAGATGATCCGGAACAGAGAGCTAGAGAGAGAGAGAGAGAG 738
 QY 181 IleValGlnGluPheGluLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
 Db 739 ATTGTACAGAGATTGGAGAGGTAGACCTTGCTGAGTGGCCGCTAGTGGCTGACTTG 798
 QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 Db 799 GAGAGAGCCCTCTTACAGATGTGTTCCTCCACCTTAACAGTCTCATCCATACAGCTTCAGAC 858
 QY 221 CysHisThrThrValArgProAlaLysProProValAlaAspArgSerLeuLysProGly 240
 Db 859 TGTCAACACTGTGAAGCCAGCTTAAGCCAGCTGTGTGACAGAGTCTTGAACCTGGA 918
 QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro 260
 Db 919 GCACCTGAGCACTCAGAAAGATATCCCAACATCGATGATGGCCGCTAGTGGTGGTCCCT 978
 QY 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValAlaGlu 280
 Db 979 GGGGGGCTGTCCCAAGTTCTCCAGTACGCCAGTCCCAACACTGCCCGGAGAGTGGAG 1038
 QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 Db 1039 ACATGTGGAATTCCTGTGGAAGAACTGATGAGAAATGATTAACATTAACCATGTTCTC 1098
 QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLysLeuPhe 320
 Db 1099 ATCCCAAGCAAGAAATGCTGGGTCTGATTACTGCAACAGAGAAAGAGAAAGAACTTTTC 1158
 QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
 Db 1159 CTCTACAGAGATCAGCAGGCTCATACACTGGGCTGGATTCTACTACCCCAACAG 1218
 QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro 360
 Db 1219 ACCCGTTTCTCTCCAGTGTGACCTACACACTCAGCTCTTAACAGATGATGTGCCA 1278
 QY 361 GluSerValAlaIleValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuThr 380
 Db 1279 GAGTCAGAGCAATGTTGCTGCTCCCAAGTTCCAGAAACTGATTTCTTAACCTACT 1338
 QY 381 AspHisGlyLeuGlnGluLysSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 Db 1339 GACCATGAGACTAGAGAGATTTCTTCCTGTGCGAGAAAGATTTCTATCCACAGAGAG 1398
 QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValIleThrIle 420
 Db 1399 GATCCACTCTGTTCTGTAGCTGACGACAGCTGACTGTGTGAGCAGAGAGAGTACCANTC 1458
 QY 421 ThrAspLeuArg 424
 Db 1459 ACAGACCTTGA 1470

RESULT 3

US-10-170-235-32565

; Sequence 32565, Application US/10170235
 ; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
 ; FILE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
 ; FILE REFERENCE: CL001380
 ; CURRENT APPLICATION NUMBER: US/10/170,235
 ; NUMBER OF SEQ ID NOS: 42514
 ; SEQ ID NO 32565
 ; LENGTH: 1767
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 ; US-10-170-235-32565

Alignment Scores:
 Pred. No.: 176-173 Length: 1767
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 46 Gaps: 0

US-09-831-452-1 (1-424) x US-10-170-235-32565 (1-1767)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
 Db 189 ATGTCTGACCATGAGATGTGAGCTCCGCCGGAAGACCGGGTACAGGCTCTCCAG 248
 QY 21 LeuGlySerAlaValAlaGluValAsnGluAspIleProPheArgArgTyrPheArgSerGly 40
 Db 249 CTGGGAGGCGCGTGAAGGTGATGATAGACATTCACCCCGGTGGTACTTCCGCTTGA 308
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGlyAsnIleGluHisAlaPhe 60
 Db 309 GTTGAGATTATCCGATGGCATTCATTACTGTGAGAGGCAACATTGAACATGCTTC 368
 QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 Db 369 ATCCCTCTAACAAAGTATATACGCTCTTATTGAGAAACCAACAAACATCGAGATTAC 428
 QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100
 Db 429 AAATCTGCTGTATCTCCGAAAAGAAAGACACAGAAAGAAATTAAGAGATTGCATT 488
 QY 101 ProTysalaglululeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
 Db 489 CCCAAGACAG 548
 QY 121 TyrAsnGluGluLysLysLysGluAlaGluGluLeuLeuLysAsnMetAlaIleGlnGln 140
 Db 549 TATATGAG 608
 QY 141 GluLeuGluLysGluLysGluLysArgValAlaGlnGlnLysGlnGlnGlnGlnGln 160
 Db 609 GAGCTGAG 668
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluLysGluLys 180
 Db 669 CAGTTCATGCTTCGAGAGAGATGATCCGGAACAGAGAGCTAGAGAGAGAGAGAGAG 728
 QY 181 IleValGlnGluPheGluLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
 Db 729 ATTGTACAGAGATTGGAGAGGTAGACCTTGCCCTAGTGGCCGCTAGTGGCTGACTTG 788
 QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 Db 789 GAGAGAGCCCTCTTGAAGTGTGTCCCACTTAACAGTCTCATCCATACAGCTTCAGAC 848
 QY 221 CysHisThrThrValArgProAlaLysProProValAlaAspArgSerLeuLysProGly 240
 Db 849 TGTCAACACTGTGAAGCCAGCTTAAGCCAGCTGTGTGACAGAGTCTTGAACCTGGA 908
 QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro 260

Db 909 GCACGTGACCACTGAAAGATATTCCCAACATCGATGGATTGGCCATGCTGTGCTCCT 968
Qy 261 G1yArgLeuCySPProGlnPheLeuGlnLeuAlaSerAlaSerThrAlaArgGlyValGlu 280
Db 969 GGGGGGCTGTCCCAAGTATTTCTCCAGTTAGCCAGTGCACACTGCCCGGGAGTGGAG 1028
Qy 281 ThrCysGlyIleLeuCySGlyLeuMetArgAsnGluPheThrIleThrHisValIleu 300
Db 1029 ACATGTGGAATTTCTCTGTGCAAAATGATGAGAAATGATTTACATTACCATTAACCATGTTCTC 1088
Qy 301 IleProlySGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
Db 1089 ATCCCCAAGCAAGAGTGGGTCTGATTACTGCAACACAGAGAAAGAAAGAACTTTTC 1148
Qy 321 LeuIleGlnAspGlnGluGlyLeuIleThrLeuGlyTyrPheIleThrHisProThrGln 340
Db 1149 CTCATACAGATTCAGCAGGGCTCTCATCACTGGAGTGGATTCTACTACCCCAACAG 1208
Qy 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro 360
Db 1209 ACCGGGTTTCTCTCCAGTGTGCACTACACACTGCTCTTACCAAGATGATGTGCCA 1268
Qy 361 GluSerValAlaIleValCysSerProlySPheGlnGluThrGlyPhePheLeuThr 380
Db 1269 GAGTCAGTAGCCATGTGTGCTCCCAAGTTCCAGGAAGTGAATCTTTAACTAACT 1328
Qy 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnGlyPheHisProHisSerTyr 400
Db 1329 GACCATGAGATGAGAGATTTCTTCTGTGCGCAGAAAGATTTCACTCCACAGCAGAG 1388
Qy 401 AspProPheLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
Db 1389 GATCCACTCTGTCTGTAGCTGACGACGACGTCGTTGGACAGAGCAGTGAACCATC 1448
Qy 421 ThrAspLeuArg 424
Db 1449 ACAGACTTCGA 1460

RESULT 4
US-09-831-452-2

; Sequence 2, Application US/09831452
; GENERAL INFORMATION:
; APPLICANT: SUGAMURA, Kazuo
; APPLICANT: TANAKA, Nobuyuki
; TITLE OF INVENTION: Protein AMSH and cDNA thereof
; FILE REFERENCE: 2001-0572A/LC/00653
; CURRENT APPLICATION NUMBER: US/09/831,452
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/JP99/06309
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: JP No. 10-322674
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1282
US-09-831-452-2

Alignment Scores:

Pred. No.: 1,87e-173 Length: 1910
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 34 Gaps: 0

US-09-831-452-1 (1-424) x US-09-831-452-2 (1-1910)

Qy 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
Db 11 ATCTTGACCACTGAGATGTGACCTCCCGCCGGAAGACCGGGTGGGGCTCTCTCCAG 70
Qy 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
Db 71 CTGGTAGTGGCGGTAGAGGTGAATGAAGACATTCACCCCGCGGATCACTCCGCTCGGA 130
Qy 41 ValGlnIleIleArgMetAlaSerIleTyrSerGluGluGlyAsnIleGlnHisAlaPhe 60
Db 131 GTTGAATTAATCCGATGGATCCATTTACTCTGAGGAAGCCACATTGAACATGCTTC 190
Qy 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluValSerProlyHisArgAspTyr 80
Db 191 ATCTCTATACAGATATACAGCTCTTATGAGAAACTACAAACATCGAGATTAC 250
Qy 81 LysSerAlaValIleProGluValSerAspThrValValLeuLeuIleAlaPhe 100
Db 251 AAATCTGCTGCTATCTCTAAAAAGAAACACACGTAAAGAAATTAAAGAGATGCTATT 310
Qy 101 ProlyAlaGluGluLeuValSerIleLeuLeuValSerTyrThrIleGluTyrThrGln 120
Db 311 CCGAAGCAGAGAGCTGAGAGCAGAGCTGTTAAACGATATACCAAGATATACAGA 370
Qy 121 TyrAsnGluGluValValSerGluValGluGluLeuAlaArgAsnMetAlaIleGln 140
Db 371 TATATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 430
Qy 141 GluLeuGluValSerGluValArgValAlaGlnGluValGlnGluGluGln 160
Db 431 GAGCTGAAAAAGAAAAAGAAAGAGGTGACAAACAAAGACAGCAATTTGAAACGGA 490
Qy 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluGluValSerGluArgLeu 180
Db 491 CAGTCCAGCTCTTCAGAGAGATGATCCGAAACAGAGCTGAAAAAGAGCACTGAA 550
Qy 181 IleValGlnGluPheGlyValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 551 ATGTACAGAGATTTGGAGAGTGAACCTGTGGCTAGGCGCTGAGCTGCTG 610
Qy 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 611 GAGAGCCCTCTTATATGTGTCTCCCACTTAACAGTCTCATCCATACAGCTTGAGAC 670
Qy 221 CysHisThrThrValArgProAlaValProValValAspArgSerLeuValProGly 240
Db 671 TGTACACAACTGTAGGCGACGCTAAGCCAGCTGTGTGACAGCGTCTTGAACCTGGA 730
Qy 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 731 GCATGAGCACTCAAGAAATTTCCCAATGATGATGCGCATGTGGTGGCT 790
Qy 261 GlyArgLeuCySPProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
Db 791 GGGCGGCTGGCCCAAGTTCTCCAGTTAGGCAAGTCCACCTCCCGGGAGTGGAG 850
Qy 281 ThrCysGlyIleLeuCySGlyLeuMetArgAsnGluPheThrIleThrHisValIleu 300
Db 851 ACATGTGGAATTTCTGTGAAAACTGATGAGAAATTAACATTCACCACTGTTCTC 910
Qy 301 IleProlySGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
Db 911 ATCCCCAAGCAAGTCTGTGCTGATTCTGCAACACAGAGAAAGAAAGAACTTTTC 970
Qy 321 LeuIleGlnAspGlnGluGlyLeuIleThrLeuGlyTyrPheIleThrHisProThrGln 340
Db 971 CTCATACAGATGAGAGGGCTCATCACTGGGTGATTCATCACTCAACCCCAACAG 1030
Qy 341 ThrAlaPheLeuSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro 360
Db 1031 ACCGGCTTCTCTCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTGCCA 1090
Qy 361 GluSerValAlaIleValCysSerProlySPheGlnGluThrGlyPhePheLeuThr 380

Db 1091 GAGTCAGTAGCATTGTTGCTCCGCCAAGTTCACAGAACTGATTCCTTAACACTACT 1150
Qy 381 ASPHISGLYLEUGLUNILLESERSEYSAARGIUNISGLYPRHEHISPROHISSEYLS 400
Db 1151 GACCATGAGCTAGAGAGATTTCTTCGTGCCAGAAAGATTTCATCCACACAGCAAG 1210
Qy 401 ASPROPOLLEUPHECYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEY 420
Db 1211 GATCACCTCTGTTCTGTAGTGTGACGACGACGACGACGACGACGACGACGAC 1270
Qy 421 ThraspleuArg 424
Db 1271 ACAGACCTTCGA 1282

RESULT 5
US-10-671-572A-2
; Sequence 2, Application US/10671572A
; GENERAL INFORMATION:
; APPLICANT: SUGAMURA, Kazuo
; APPLICANT: TANAKA, Nobuyuki
; TITLE OF INVENTION: Protein AMSH and cDNA thereof
; FILE REFERENCE: 2003-1382/MMC/00653
; CURRENT APPLICATION NUMBER: US/10/671,572A
; PRIORITY FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: 09/831,452
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/JP99/06309
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: JP No. 10-322674
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1910
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1282
; US-10-671-572A-2

Alignment Scores:
Pred. No.: 1,87e-173 Length: 1910
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-831-452-1 (1-424) x US-10-671-572A-2 (1-1910)

Qy 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
Db 11 ATGCTGACCATGAGATGTGAGCTCCGCCAGAAACCGGAGAGGCTCTCTCCAG 70
Qy 21 LeuGlySerAlaValGluValAsnGluAspLeuProArgArgIyrPheAspSerGly 40
Db 71 CTGGGTAGTGGCGTAGAGGTGAAGAGACATTCACCCCGTGGTACTCCGCTCGAG 130
Qy 41 ValGluLeuLeuAspMetAlaSerLeuIyrSerGluGluGluValAsnLeuGluAlaPhe 60
Db 131 GTTAGAGATTATCCAAAGGATCCATTCATCTGAGAGAGCAACATTGAACATGCTTC 190
Qy 61 IleuLeuIyrAsnLeuIyrLeuPheLeuPheLeuGluLeuProLysHisAspAspIyr 80
Db 191 ATCTCTATTAACAAGTATATACGCTCTTATTTGAGAACTACCAAAACATCCAGATTAC 250
Qy 81 LysSerAlaValIleProGluValLysAspThrValLysLysLeuLysGluIleAlaPhe 100
Db 251 AAATCTGCTGCTGCTGAAAGAAAGAACACAGATAAGAAATTAAAGAGATTGCAATT 310
Qy 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgIyrThrLysGluIyrThrGlu 120

Db 311 CCCAAAGCAGAGAGCTGAGAGAGCTGTTAAACATATATACCAAGAAATATACAGAA 370
Qy 121 TyrAsnGluLysLysLysGluValGluGluLeuValAspAspMetAlaIleGlnGln 140
Db 371 TATTAAGAAAGAAAG 430
Qy 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnLysGlnGln 160
Db 431 GAGCTGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 490
Qy 161 GlnPheHisAlaPheGluLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 180
Db 491 CAGTTCCAGCTTCGAGAGAGATGATCCGAGACCGAGAGAGAGAGAGAGAGAGAG 550
Qy 181 IleValGlnGluPheGlyLysValAspProGluLeuGlyGlyProLeuValProAspLeu 200
Db 551 ATTGTACAGAGATTGGAGAGGTACACCTGGCTAGGTGGCCCGCTAGTGGCTAGCT 610
Qy 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerLeuGlnProSerAsp 220
Db 611 GAGAGAGCTCTCTTGATGTGTCCCACTTAAACAGTCTATCATCATCACCTTCAGAC 670
Qy 221 CysHisThrThrValArgProAlaLysProProValValAspAspSerLeuProGly 240
Db 671 TGTCAACACTGTATAGGACCACTAGCCACTGTGTGACAGCTCTTAAACCTGCA 730
Qy 241 AlaLeuSerAsnSerGluSerLeuProThrIleAspGlyLeuArgHisValValPro 260
Db 731 GCACAGAGCACTCAGAAAGATTTCCCAATTCAGATGATTCGCGCATGTGGTGCT 790
Qy 261 GlyArgLeuCysProGluPheLeuGluLeuValAspAlaAspThrAlaArgGlyValGlu 280
Db 791 GGGCGGCTGTGCCACAGTTCTCCAGTTAGCGAGTGCACCAACCTGCGGGAGGTGAG 850
Qy 281 ThrCysGlyLeuLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
Db 851 ACATGTGAGATTCTGTGAGAAACTGATGAGATGATTTACATTAACCATTCATGTTCTC 910
Qy 301 IleProLysGlnSerAlaGlySerAspIyrCysAsnThrGluAsnGluGluLeuPhe 320
Db 911 ATCCCAAGCAAGAGCTGGGTGTGATTCGCAACACAGAGAGAGAGAGAGAGAGAG 970
Qy 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyIyrPheHisThrHisProThrGln 340
Db 971 CTCATACAGATCAGAGAGGCTCATACACTGGGCTGGATTATACACCCACACAG 1030
Qy 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerIyrGlnMetLeuPro 360
Db 1031 ACCGCGTTCTCTCCAGTGTGACCTACACACTGCTCTTACAGATGATGTGCA 1090
Qy 361 GluSerValAlaIleValCysSerProLysPheGlnGlnIyrThrGlyPhePheLeuThr 380
Db 1091 GAGTCAGTAGCATTGTTGCTCCGCCAAGTTCAGAACTGAGATTCTTAAACTACT 1150
Qy 381 ASPHISGLYLEUGLUNILLESERSEYSAARGIUNISGLYPRHEHISPROHISSEYLS 400
Db 1151 GACCATGAGCTAGAGAGATTTCTTCGTGCCAGAAAGATTTCATCCACACAGCAAG 1210
Qy 401 ASPROPOLLEUPHECYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEYSEY 420
Db 1211 GATCACCTCTGTTCTGTAGTGTGACGACGACGACGACGACGACGACGACGAC 1270
Qy 421 ThraspleuArg 424
Db 1271 ACAGACCTTCGA 1282

RESULT 6
US-09-770-175-8814
; Sequence 8814, Application US/09770175
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 1600,2058-001
 ; CURRENT APPLICATION NUMBER: US/09/770,175
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: US 60/178,874
 ; PRIOR FILING DATE: 2000-01-28
 ; NUMBER OF SEQ ID NOS: 8967
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8814
 ; LENGTH: 2001
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-770-175-8814

Alignment Scores:
 Pred. No.: 1,98e-173 Length: 2001
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-831-452-1 (1-424) x US-09-770-175-8814 (1-2001)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
 Db 66 ATGCTGACCATGAGATGTGAGCGCTCCGCCGAAACCGGGGAGGCGCTCTCCACAG 125
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 Db 126 CTGGGTGTGGCGTAGAGGTGAATGAAGACATTCACCCCGGTGGTACTCCGCTCGGA 185
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluValMetIleGluAlaPhe 60
 Db 186 GTTGAGATATCCGAAATGGCATCTTACTGAGGAAAGCAACATGTGAACATGCGCTTC 245
 QY 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluIleLeuProGluIleAspTyr 80
 Db 246 ATCTCTATTAACAAGATATATACGCTCTTTATTAGAAACATCAAAACATCGAGATTAC 305
 QY 81 LysSerAlaValIleProGluIleAspThrValIleIleIleIleIleIleIleIleIle 100
 Db 306 AATCTGCTGCTATCTCCGAAAGAAAGACACAGTAAGAAATTAAGAGATTGCACTT 365
 QY 101 ProLysAlaGluGluLeuLeuValGluLeuLeuIleValGluLeuIleValGluLeu 120
 Db 366 CCCAAAGCAAGAGCTGAAAGGAGAGCTGTAAAGGATATACCAAGAAATATACAGAA 425
 QY 121 TyrAsnGluGluIleValGluValGluAlaGluIleValIleValIleValIleValIle 140
 Db 426 TATTAATGAAGAAAG 485
 QY 141 GluLeuGluIleValGluValGluValGluValGluValGluValGluValGluValGlu 160
 Db 486 GAGCTGGAAG 545
 QY 161 GlnPheHisAlaPheGluGluIleValGluValGluValGluValGluValGluValGlu 180
 Db 546 CAGTTCCATGCTTCCGAGAGATATCCGAAACAGAGAGAGAGAGAGAGAGAGAGAGAG 605
 QY 181 IleValGlnGluPheGluValGluValGluValGluValGluValGluValGluValGlu 200
 Db 606 ATTGAACAGAGATTGGGAAGTATGACCTGCTAGAGAGAGAGAGAGAGAGAGAGAGAG 665
 QY 201 GluLysProSerLeuAspValIlePheProThrIleThrValIleSerSerIleGlnProSer 220
 Db 666 GAGAGAGCCCTCTTATGATGTGTCTCCCACTTAACAGCTCATATACAGCTTCAGAC 725
 QY 221 CysHsThrThrValAlaProAlaIleAspProProValAlaAspArgSerLeuLysProGly 240
 Db 726 TGTCAACAACCTGTAGAGCCAGCTAAGCCACCTGTGGTGGAGAGAGAGAGAGAGAGAG 785

QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro 260
 Db 786 GCATGAGCAACCTCAGAAAGATATCCCAACATCGATGTGCGCATGTGTGTGCT 845
 QY 261 GlyArgLeuCysProGluPheLeuGluLeuIleAsnAlaAsnThrAlaArgGlyValGlu 280
 Db 846 GGGCGGCTGTGCCACAGATTTCTCCAGTTAGCAAGTCCCAACACTGCCGGGAGGTGAG 905
 QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 Db 906 ACATGTGAAATCTGTGGAAGAACTGATGAGCATGAAATTTACCATTAACCATGTTC 965
 QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 Db 966 ATCCCAAGCAAGAGCTGGGTCTGATTATCGCAACAGAGAAAGCAAGAACTTTTC 1025
 QY 321 LeuIleGluAspGluGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
 Db 1026 CTCAATACAGATCAGCAGAGGCTCATACACTGGGCTGGATTCAATCAACCCACAG 1085
 QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGluMetLeuPro 360
 Db 1086 ACCGCGTTCTCTCCAGTGTGACCTACACTCAGCTGCTTACAGATGATGTGCA 1145
 QY 361 GluSerValAlaIleValCysSerProLysPheGlnIleThrGlyPhePheIleThr 380
 Db 1146 GAGTCAGTGGCATTTGCTCTCCCAAGTTCCAGAAACTGGAATCTTTAACTACT 1205
 QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnGlyPheHisProHisSerLys 400
 Db 1206 GACCATGAGATGAGAGATTTCTTCCTGTCGCGCAAGAAAGATTCATCAACAGCAAG 1265
 QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
 Db 1266 GATCCACCTCTGTTCTGTGATGAGCAGCAGGACTGTGTGAGAGAGAGAGAGAG 1325
 QY 421 ThrAspLeuArg 424
 Db 1326 ACAGACCTTCCA 1337

RESULT 7
 US-10-752-442-1
 ; Sequence 1, Application US/10752442
 ; GENERAL INFORMATION:
 ; APPLICANT: Itoh, Fumiko
 ; APPLICANT: Itoh, Susumu
 ; APPLICANT: Heldin, Carl-Henrik
 ; APPLICANT: ten-Dijke, Peter
 ; TITLE OF INVENTION: SMAD ASSOCIATING POLYPEPTIDES
 ; FILE REFERENCE: L0461,70096US01
 ; CURRENT APPLICATION NUMBER: US/10/752,442
 ; CURRENT FILING DATE: 2004-01-06
 ; PRIOR APPLICATION NUMBER: US 09/665,479
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US 60/154,846
 ; PRIOR FILING DATE: 1999-09-20
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 2003
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (63)..(1334)
 US-10-752-442-1

Alignment Scores:
 Pred. No.: 1,98e-173 Length: 2003
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0

DB: 55 Gaps: 0

US-09-831-452-1 (1-424) x US-10-752-442-1 (1-2003)

QY 1 MetSerSPHISGIYASPVASerLeuProProGluAspArgValArgAlaLeuSerGln 20

DB 63 ATGTCTGACCATGGAGATGTGAGCTCCCGCCGAAACCGGGTGAAGGCTCTCTCCGAG 122

QY 21 LeuGIYSerAlaValAGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40

DB 123 CTGGGTAGTGGGTAGAGGTGAATGAAGACATTCACCCCGTGGTACTTCCCTCTGGA 182

QY 41 ValGIuIleIleArgMetAlaSerIleTyrSerGluGluGluYasnIleGluHisAlaPhe 60

DB 183 GTTAGATTAATCCGAATGCGATTCATTACTCTGAGAAAGCAACATTAACATGCTTC 242

QY 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluIleuLeuProLysHisArgAspTyr 80

DB 243 ATCTCTTAATAACAAGTATATCACCGCTCTTATTGAGAACTACCAAAACATTCAGATTAAC 302

QY 81 LysSerAlaValIleProGluIleuLysAspThrValIleLysIleuLysGluIleAlaPhe 100

DB 303 AAATCTGCTGATCTCTGAAAAGAAAGACAGTAAAGAAATTAAAGAGATTGCACTTT 362

QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu 120

DB 363 CCCAAAGCAGAGAGCTGAAGGCGAGAGCTGTAAAGATATACCAAAAGATATACAGAA 422

QY 121 TyrAsnGluGluLysLysGluAlaGluGluLeuAlaArgSerMetAlaIleGlnGln 140

DB 423 TATATGAGAGAAAGAAAGAAAGAAAGAGAGAAATGCGCGAAGACAGGCTACAGCA 482

QY 141 GluLeuGluLysGluLysArgValAlaGlnGlnLysGlnGlnGlnGlnGlnGln 160

DB 483 GAGCTGGAAGAAAGAAAGAAAGAGGTAGCAACAAGAGCAGAGCATTTGGAACAGGAA 542

QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluGluLysGluLysGluLys 180

DB 543 CAGTTCCATGCTTCGAGAGATGATCCGAGACAGAGCTTAAGAAAGCGACATGAA 602

QY 181 IleValGlnGluPheGluLysValAspProGluYleuGluYleProLeuValProAspLeu 200

DB 603 ATTGTACAGAGATTTGGGAAGTGAAGCTCGCTGAGTGGCTGAGCTGACCTGACTGG 662

QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220

DB 663 GAGAAAGCCCTCTTACATGTGTTCCCACTTAACAGTCTCATCCATACAGCTTCAGAC 722

QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGlu 240

DB 723 TGTACACAACTGTAAAGCGCACTTAAGCCACTGTGTGACAGGTCTTGAACCTGGA 782

QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGluLeuArgHisValValValPro 260

DB 783 GCACTGAGCAACTCAGAAAGATATCCCAATCATGATTTGCGCCATGTGGTGGCT 842

QY 261 GluYArgLeuLysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGluValGlu 280

DB 843 GGGGGGTGTGGCCACAGTTCTCCAGTTAGCCAGTGCACACCTCCCGGGAGTGGAG 902

QY 281 ThrCysGluIleLeuLysGluLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300

DB 903 ACAATGTGAATTTCTCTGAAAACATGATGAGGAATTAATTCATTCACCACTGTTCC 962

QY 301 IleProLysGlnSerAlaGluLysArgTyrCysAsnThrGluAsnGluGluLeuPhe 320

DB 963 ATCCCAAGCAAAATGTGGGTCTGATTTACTGCAACAGAGAAACAAAGAAATTTTC 1022

QY 321 LeuIleGlnAspGlnGlnGluLeuIleThrLeuGluTyrIleHisThrHisProThrGln 340

DB 1023 CTCTATAAGATTCAGCAGGCGCTTATACATGCGGTGATTCATCTACCCCAACAG 1082

QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360

DB 1083 ACCGGTTTCTCTCAGTGTGACCTTACACTACCTCTTACAGATATGTGGCA 1142

QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLeuThr 380

DB 1143 GAGTCAGTAGGCATGTGTGCTCCCGCAAGTTCACAGAACTGATTTCTTAACTAACT 1202

QY 381 AspHisGluLeuGluIleLeuSerCysArgGlnLysGlyPheHisProHisSerLys 400

DB 1203 GACCAAGACTAGAGAGATTTCTTCTGTCGCGCAAGAAAGATTTCACTCAACAGCAG 1262

QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420

DB 1263 GATCCACTCTGTCTGTAGCTGACGACGCTGATGTGGACAGAGCAGTGAACATC 1322

QY 421 ThrAspLeuArg 424

DB 1323 ACAGACTTCGA 1334

RESULT 8

US-09-881-797-3931/C

Sequence 3931, Application US/09881797

GENERAL INFORMATION:

APPLICANT: Tim Keith

TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES

TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY

FILE REFERENCE: HMO01-04

CURRENT APPLICATION NUMBER: US/09/881,797

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: 60/211,749

PRIOR FILING DATE: 2000-06-14

NUMBER OF SEQ ID NOS: 4687

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3931

LENGTH: 2045

TYPE: DNA

ORGANISM: Human

US-09-881-797-3931

Alignment Scores:

Pred. No.: 2,046-173 Length: 2045

Score: 2208.00 Matches: 424

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 36 Gaps: 0

US-09-831-452-1 (1-424) x US-09-881-797-3931 (1-2045)

QY 1 MetSerSPHISGIYASPVASerLeuProProGluAspArgValArgAlaLeuSerGln 20

DB 1929 ATGTCTGACCATGGAGATGTGAGCTCCCGCCGAAACCGGGTGAAGGCTCTCTCCGAG 1870

QY 21 LeuGIYSerAlaValAGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40

DB 1869 CTGGGTAGTGGCGTGAAGAGTGAATGAAGACATTCACCCCGTGGTACTTCCGCTGGA 1810

QY 41 ValGIuIleIleArgMetAlaSerIleTyrSerGluGluGluYasnIleGluHisAlaPhe 60

DB 1809 GTTGAGATTATCCGAATGCGATTCATTACTCAGAGAGGCAACATTGAACATGCCCTTC 1750

QY 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluIleuLeuProLysHisArgAspTyr 80

DB 1749 ATCTCTTAATAACAAGTATATCACCGCTCTTATTGAGAAACATCAACAAACATCGATTAC 1690

QY 81 LysSerAlaValIleProGluLysLysAspThrValIleLysIleuLysGluIleAlaPhe 100

DB 1689 AAATCTGCTGATCTCTGAAAAGAAAGACAGTAAAGAAATTAAGAGGATTTGCACTTT 1650

QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu 120

DB 1629 CCCAAAGCAGAGAGCTGAAGGCGAGAGCTGTAAAGCATATACCAAGAAATATACAGAA 1570

QY 121 TyrSnglGluGlySlySglValAlaGluGluLeuValaArgaMetAlaIleGlnGln 140
Db 1569 TATATGAGAAAGAAAGAAAGAGAGAGATTTGGCCCGGAACTGGCCATCCACAA 1510
QY 141 GluGluGlySglGlySglValAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 1509 GAGCTGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1450
QY 161 GluPheHisAlaPheGluGluMetIleArgSnglGlnGluGlySglValArgLeuLys 180
Db 1449 CAGTTCCATGCTTCGAGAGAGATGATCCGGAACAGAGAGAGAGAGAGAGAGAG 1390
QY 181 IleValGlnGluPheGlyValValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 1389 ATTGTACAGAGATTGGAGAGATGAGCCCTGGCTAGTGCCCTAGTGCCCTAG 1330
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 1329 GAGAAAGCCCTCTTATGATGTTTCCCACTTAACAGTCTCATTCATACAGCTTCAGAC 1270
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
Db 1269 TGTACACAACTGTAAAGCCAGCTAAGCCAGCTGTGTGAGCAGGTCTTGAAACCTGGA 1210
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 1209 GCACGTGAGCACTAGAAAGATTTCCCAATCATGATGATTTGCCCATGTGTGTGCT 1150
QY 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
Db 1149 GGGGGCTGTGCCACAGTTCTCCAGTTAGCCAGTCCCAACAATGCCCGGAGGTGAG 1090
QY 281 ThrCysGlyIleLeuCysGlyValLeuMetArgSnglPheThrIleThrHisValLeu 300
Db 1089 ACAAGTGAATCTCTGTGAGAAAGCTGATGAGAAATGATTAACATTAACCAATGTTCTC 1030
QY 301 IleProLysGlnSerIleGlySerAspLysCysAsnThrGluAsnGluGluLeuPhe 320
Db 1029 ATCCCCAGCAAGTGTCTGATTCGATTCGCAACAGAGAACAGAAAGAACTTTTC 970
QY 321 LeuIleGlnAspGlnGlyLeuIleThrLeuGlyIleThrIleHisIleProThrGln 340
Db 969 CTCATACAGATCAGAGAGCCCTCATCACTGGCTGATTCATCACTCCACCCACAG 910
QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerIleArgIleMetLeuPro 360
Db 909 ACCCGTTCCTCCAGTGTGAGCTACACACTGCTCTTAACAGATGATTTGCCA 850
QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrArgIlePhePheLysLeuThr 380
Db 849 GAGTCAGTAGCATTTGTTGCTCCCCCAAGTTCCAGAGAACTGATTTTAACTAACT 790
QY 381 AspiSglLeuGluGluIleSerSerCysArgGlnGlyPheHisProHisSerLys 400
Db 789 GACCATGAGATGAGAGATTTCTTCCTGTCAGAGAAAGATTTCTATCCACAGCAG 730
QY 401 AspProProLeuPheCysSerCysSerHisValThrValValAspArgAlaValThrIle 420
Db 729 GATCCACTCTGTCTGTAGCTGAGCCAGCTGATGCTGTGTGAGCAGAGATGACCATC 670
QY 421 ThrAspLeuArg 424
Db 669 ACAGACCTTCA 658

RESULT 9
US-10-021-698-3931/c
; Sequence 3931, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: HMO1-04
; CURRENT APPLICATION NUMBER: US/10/021,698

QY 121 TyrSnglGluGlySlySglValAlaGluGluLeuValaArgaMetAlaIleGlnGln 140
Db 1569 TATATGAGAAAGAAAGAAAGAGAGAGATTTGGCCCGGAACTGGCCATCCACAA 1510
QY 141 GluGluGlySglGlySglValAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 1509 GAGCTGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1450
QY 161 GluPheHisAlaPheGluGluMetIleArgSnglGlnGluGlySglValArgLeuLys 180
Db 1449 CAGTTCCATGCTTCGAGAGAGATGATCCGGAACAGAGAGAGAGAGAGAGAGAG 1390
QY 181 IleValGlnGluPheGlyValValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 1389 ATTGTACAGAGATTGGAGAGATGAGCCCTGGCTAGTGCCCTAGTGCCCTAG 1330
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 1329 GAGAAAGCCCTCTTATGATGTTTCCCACTTAACAGTCTCATTCATACAGCTTCAGAC 1270
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
Db 1269 TGTACACAACTGTAAAGCCAGCTAAGCCAGCTGTGTGAGCAGGTCTTGAAACCTGGA 1210
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 1209 GCACGTGAGCACTAGAAAGATTTCCCAATCATGATGATTTGCCCATGTGTGTGCT 1150

QY 121 TyrSnglGluGlySlySglValAlaGluGluLeuValaArgaMetAlaIleGlnGln 140
Db 1569 TATATGAGAAAGAAAGAAAGAGAGAGATTTGGCCCGGAACTGGCCATCCACAA 1510
QY 141 GluGluGlySglGlySglValAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 1509 GAGCTGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1450
QY 161 GluPheHisAlaPheGluGluMetIleArgSnglGlnGluGlySglValArgLeuLys 180
Db 1449 CAGTTCCATGCTTCGAGAGAGATGATCCGGAACAGAGAGAGAGAGAGAGAGAG 1390
QY 181 IleValGlnGluPheGlyValValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 1389 ATTGTACAGAGATTGGAGAGATGAGCCCTGGCTAGTGCCCTAGTGCCCTAG 1330
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 1329 GAGAAAGCCCTCTTATGATGTTTCCCACTTAACAGTCTCATTCATACAGCTTCAGAC 1270
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
Db 1269 TGTACACAACTGTAAAGCCAGCTAAGCCAGCTGTGTGAGCAGGTCTTGAAACCTGGA 1210
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 1209 GCACGTGAGCACTAGAAAGATTTCCCAATCATGATGATTTGCCCATGTGTGTGCT 1150

Alignment Scores:
Pred. No.: 2,046-173 Length: 2045
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 43 Gaps: 0

US-09-831-452-1 (1-424) x US-10-021-698-3931 (1-2045)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValaArgAlaLeuSerGln 20
Db 1929 ATTCTGACCACTGAGATGTGAGCTCCCGCCGAAAGACGGGTGAGGGCTCTTCCAG 1870
QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgIleThrArgSerGly 40
Db 1869 CTGGGTAGTGCCTGATGAGTGAATGAAACATTCACCCCGGTGATCTTCCGCTTGA 1810
QY 41 ValGluIleIleArgMetAlaSerIleIleIleSerGluGluGlyAsnIleGluHisAlaPhe 60
Db 1809 GTTGAGATTTATCCGATGAGCATTCATTTCTCTGAGAGAGCAACATTAACATGCTTC 1750
QY 61 IleLeuThrAsnLysTyrlleThrLeuPheIleGluLysLeuProLysHisArgAspTyrl 80
Db 1749 ATCCTTATTAACAGATATATCAAGCTCTTATTTGAGAAACATCAACAAATGAGATTAC 1690
QY 81 LysSerAlaValIleProGluLysAspThrValLysLeuLysLeuGluIleAlaPhe 100
Db 1689 AAATGCTGTCACTTCTTAAAGAAAGACAGAGAAATTAAGAGAGATTTGCATTT 1630
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrlleThrLysGluTyrlleGln 120
Db 1629 CCCAAAGCAGAAAGAGTGAAGGAGAGCTGTTAAACATATATACCAAGATATACAGAA 1570
QY 121 TyrSnglGluGlySlySglValAlaGluGluLeuValaArgaMetAlaIleGlnGln 140
Db 1569 TATATGAGAAAGAAAGAAAGAGAGAGATTTGGCCCGGAACTGGCCATCCACAA 1510
QY 141 GluLeuGluLysGlnArgValAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Db 1509 GAGCTGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1450
QY 161 GluPheHisAlaPheGluGluMetIleArgSnglGlnGluGlySglValArgLeuLys 180
Db 1449 CAGTTCCATGCTTCGAGAGAGATGATCCGGAACAGAGAGAGAGAGAGAGAGAG 1390
QY 181 IleValGlnGluPheGlyValValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 1389 ATTGTACAGAGATTGGAGAGATGAGCCCTGGCTAGTGCCCTAGTGCCCTAG 1330
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 1329 GAGAAAGCCCTCTTATGATGTTTCCCACTTAACAGTCTCATTCATACAGCTTCAGAC 1270
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
Db 1269 TGTACACAACTGTAAAGCCAGCTAAGCCAGCTGTGTGAGCAGGTCTTGAAACCTGGA 1210
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 1209 GCACGTGAGCACTAGAAAGATTTCCCAATCATGATGATTTGCCCATGTGTGTGCT 1150

QY 261 GlyArgLeuCySPROGlnPheLeuGlnLeuAlaSerLysThrAlaArgGlyValGlu 280
 Db 1149 GGGGGGCTGGTCCACAGTTTCTCCAGTTAGCCAGTCCACACTGCCCCGGAGTGGAG 1090
 QY 281 ThrCysGlyIleLeuCySGlyLysLeuMetArgAsnGlnPheThrIleThrHisValIleu 300
 Db 1089 ACATGTGAATTCCTCTGTGAAACCTGATGAGATGAAATTTACCATTTACCATTTCTC 1030
 QY 301 IleProlySGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 Db 1029 ATCCCAAGCAAGTGTGGTCTGATTACTGCAACAGAGAAACGAAAGAACTTTTC 970
 QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyIlePheIleThrHisProThrGln 340
 Db 969 CTCATACGATCGACAGGAGCTCATCACTGGGCTGGATTCATCACTCAACCCACAGAG 910
 QY 341 ThrIlePheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360
 Db 909 ACCCGTTTCTCTCTCAGTGTGACCTACACACTCACTGCTTAAACGATGATTTGCCA 850
 QY 361 GluSerValAlaIleValCysSerProlySPheGlnGluThrGlyPhePheLysLeuThr 380
 Db 849 GAGTCAGAGCAATGTTGTCTCCCAAGTCCAGAGAACTGGATTCTTTAAACTTAAT 790
 QY 381 AspHisGlyLeuGlnGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 Db 789 GACCATGAGACTAGAGAGATTTCTTCTGTGCGCAGAAAGATTTATCCACACAGAGAG 730
 QY 401 AspProPheLeuPheCysSerCysSerHisValThrValValAspArgAlaValThrIle 420
 Db 729 GATCCACCTCTGTTCTGTAGCTGACGACGATGACTGTGTGTGACAGAGCAGTACACT 670
 QY 421 ThrAspLeuArg 424
 Db 669 ACAGACCTTCCA 658

RESULT 10

US-60-213-362-4025
 : Sequence 4025, Application US/60213362
 : GENERAL INFORMATION:
 : APPLICANT: Morris, MacDonald
 : APPLICANT: Lal, Preeti
 : APPLICANT: Deep, Dink
 : TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
 : TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
 : FILE REFERENCE: GX-0016 P
 : CURRENT APPLICATION NUMBER: US/60/213,362
 : NUMBER OF FILING DATE: 2000-06-21
 : NUMBER OF SEQ ID NOS: 8429
 : SOFTWARE: PERL Program
 : SEQ ID NO 4025
 : LENGTH: 2087
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc feature
 : OTHER INFORMATION: Incyte ID No: 235222.9
 US-60-213-362-4025

Alignment Scores:

Prod. No.: 2,098-173 Length: 2087
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 77 Gaps: 0

US-09-831-452-1 (1-424) x US-60-213-362-4025 (1-2087)

QY 1 MetSerAspHisGlyAspValSerLeuProGlnAspArgValArgAlaLeuSerGln 20
 Db 189 ATGTGTACCATGAGATGTGAGCTCCCGCCAGAACCGGGGTGAGGGCTTCTCTCCAG 248

QY 21 LeuGlySerAlaValGluValAsnGluAspIleProPheArgArgTyrPheArgSerGly 40
 Db 249 CTGGGTATGTGGGTAGAGGTAAATGAAAGATTCACCCCTGGTACTTCGGCTCTGGA 308
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGlnGluGlyAsnIleGluHisAlaPhe 60
 Db 309 GTTGAATTAATCCGAATGCGATTCATTTACTCTGAGAGAGGCAACATTTGAACATGCCCTC 368
 QY 61 IleTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 Db 369 ATCCCTATTAACAGATATATACAGCTCTTTTATGAAACCTACCAAACTAGGAGTTAC 428
 QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100
 Db 429 AATCTGCTGTCTATTCCTGAAAGAAAGACACAGTAAAGAAATTTAAAGAGATTCATTT 488
 QY 101 ProlySalaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGlyTyrThrGlu 120
 Db 489 CCAAG 548
 QY 121 TyrAsnGluGluLysLysGluAlaGluGluLeuIleAspAsnMetAlaIleGlnGln 140
 Db 549 TATTAATGAGAAAGAAAG 608
 QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnLeuGluGlnGlu 160
 Db 609 GACGTGAGAAAG 668
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluLysLys 180
 Db 669 CAGTTCATGCTTCGAGAGATGATTCGAGAACAGAGAGAGAGAGAGAGAGAGAGAGAG 728
 QY 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
 Db 729 ATTGTACAGAGATTTGGAGAGTACACCTGCTAGGTGGCCCGCTAGTCCGACTTG 788
 QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 Db 789 GAG 848
 QY 221 CysHisThrThrValArgProAlaLysProProValValAspAspSerLeuValProGly 240
 Db 849 TGTCAACAGACTGTAAGGAG 908
 QY 241 AlaLeuSerAsnSerGlnSerIleProThrIleAspGlyLeuArgHisValValValPro 260
 Db 909 GCACTGAGACCACTGAGAAAGATTTCCCAATTCATGATTTGCCCATGTGTGTGCT 968
 QY 261 GlyArgLeuCySPROGlnPheLeuGlnLeuAlaSerLysThrAlaArgGlyValGlu 280
 Db 969 GGGGGGCTGGTCCACAGTTTCTCCAGTTAGCCAGTCCACACTGCCCCGGAGTGGAG 1028
 QY 281 ThrCysGlyIleLeuCySGlyLysLeuMetArgAsnGlnPheThrIleThrHisValIleu 300
 Db 1029 ACATGTGAATTCCTCTGTGAAACCTGATGAGATGAAATTTACCATTTACCATTTCTC 1088
 QY 301 IleProlySGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 Db 1089 ATCCCAAGCAAGTGTGGTCTGATTACTGCAACAGAGAAACGAAAGAACTTTTC 1148
 QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyIlePheIleThrHisProThrGln 340
 Db 1149 CTCATACGATCGACAGGAGCTCATCACTGGGCTGGATTCATCACTCAACCCACAGAG 1208
 QY 341 ThrIlePheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360
 Db 1209 ACCCGTTTCTCTCTCAGTGTGACCTACACACTCACTGCTTAAACGATGATTTGCCA 1268
 QY 361 GluSerValAlaIleValCysSerProlySPheGlnGluThrGlyPhePheLysLeuThr 380
 Db 1269 GAGTCAGAGCAATGTTGTCTCCCAAGTCCAGAGAACTGGATTCTTTAAACTTAAT 1328

```
QY 381 AsphHisGlyLeuGluGluLeuSerSerCySarGlnGlyPheHisProHisSerLys 400
Db 1329 GACCATGACATGAGAGAGATTCTTCCTGCGCAGAAAGATTTCACACAGCAAG 1388
QY 401 AspProPheLeuPheCySerCySerHisValThrValAlaAspAlaValThrIle 420
Db 1389 GATCACCTCTGTTCTGTAGTGCAGCAGCTGATGCTGTGTGACAGAGCATGACCATC 1448
QY 421 ThrAspLeuArg 424
Db 1449 ACAGACCTTCGA 1460

RESULT 11
US-60-278-561-6951
; Sequence 6951, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0012-1 P
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 6951
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 235222.9
US-60-278-561-6951

Alignment Scores:
Pred. No.: 2,098-173 Length: 2087
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 84 Gaps: 0

US-09-831-452-1 (1-424) x US-60-278-561-6951 (1-2087)
QY 1 MetSerAspHisGlyAspValSerLeuProGlnAlaAspValArgAlaLeuSerGln 20
Db 189 ATGTCTGACCATGAGATGATGAGCTCCCGCCGAGAACCGGGTGAAGGCTCTTCCACG 248
QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgIlePheArgSerGly 40
Db 249 CTGGGTAGTGGGTAGAGAGTGAATGAGACATTCACCCCGTGGTACTTCCGCTCGGA 308
QY 41 ValGluIleLeuArgMetAlaSerIleIleIleIleIleIleIleIleIleIleIle 60
Db 309 GTTAGATTATCCGAATGCAATCCATTACTCTGAGAGAGCAACATTGAACATGCCCTTC 368
QY 61 IleLeuIleAsnLysIleIleThrLeuPheIleGluIleLysLeuProLysHisArgAspTyr 80
Db 369 ATCCCTGATTAACAAGTATATCAGCTCTTTATTTAGAAACATCAACAAACATCGAGATTAC 428
QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100
Db 429 AAATCTGCTGTCATCTCCGAAAAAGAACACAGTAAAGAAATTAAGAGGATTGCATTT 488
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
Db 489 CCCAAAGAGAGAGCTGAAGCGAGCTGTTAAACGATATACCAAGATATACGAA 548
QY 121 TyrAsnGluGluLysLysGluAlaGluGluLeuLysArgAsnMetAlaIleGlnGln 140
Db 549 TATATGAAGAAAGAGAGAGAGAGCAAGAAATTGGCCCGAACAATGCGCATTCACAGAA 608
```

```
QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnLeuGluGlnGlu 160
Db 609 GAGCTGGAAAAAGAAAAACAGAGGTAGACAAACAGAACAGACAGCAATTTGGAAACGGAA 668
QY 161 GluPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgLeuLys 180
Db 669 CAGTTCCATGCTCTTCAGAGATGATCCGGAACAGAGACTAGAAAAAGAGCACTGAAA 728
QY 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
Db 729 ATTGTACAGAGATTGGAGAGTAGACCTGCGCTAGTGAGCCCGCTAGTGCCTGACTTG 788
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerArg 220
Db 789 GAGAACCCCTCTTATGATGTGTTCCACCTTAACTGATCATCATACAGCCTTCAGAC 848
QY 221 CysHisThrThrValArgProAlaLysProProValAlaAspArgSerLeuLysProGly 240
Db 849 TGTCAACAACACTGTAAGCGCAGCTAGCACCCTGTGTGAGACAGTCTTGAACCTGGA 908
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
Db 909 GCACGTAGCAACTCAGAAAAAGTATCCCAACATCGATGATGCGCATGTGTGGTGCCT 968
QY 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
Db 969 GAGCGGCTGTGCCACAGATTCTTCCAGTTAGCCAGTGCACACATGCGGGAGTGGAG 1028
QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
Db 1029 ACATGTGGAATCTCTGTGGAACACTGATGAGAAATGAATTACCATTCACCATGTTCTC 1088
QY 301 IleProLysGlnSerIleAlaLysSerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
Db 1089 ATCCCAAGCAAAAGTCTGGGTCTGATTACTGCACACAGAAACGAAAGAACTTTTC 1148
QY 321 LeuIleGlnAspGlnGlnLysLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
Db 1149 CTATACAGAGATCAGAGGGCCCTCATCACATCGGGTGTGATTCATACCTCACCCACAG 1208
QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360
Db 1209 ACCGCTTTCTCTCCAGTGTGACACACTCATCTCTTACCAGATGATGTGGCA 1268
QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGluPhePheLysLeuThr 380
Db 1269 GAGTCAGTAGCCATTTGTTGCTCCCCAGATTCCAGGAACCTGATCTTTAAACTTACT 1328
QY 381 AsphHisGlyLeuGluLysLeuSerSerCySarGlnLysGlyPheHisProHisSerLys 400
Db 1329 GACCATGACATGAGAGATTCTTCCTGCGCAGAAAGATTTCACACAGCAAG 1388
QY 401 AspProPheLeuPheCySerCySerHisValThrValAlaAspAlaValThrIle 420
Db 1389 GATCACCTCTGTTCTGTAGTGCAGCAGCTGATGCTGTGTGACAGAGCATGACCATC 1448
QY 421 ThrAspLeuArg 424
Db 1449 ACAGACCTTCGA 1460

RESULT 12
US-60-324-185-16983
; Sequence 16983, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
```

CURRENT FILING DATE: 2001-09-21
 NUMBER OF SEQ ID NOS: 35862
 SOFTWARE: PERL Program
 SEQ ID NO 16983
 LENGTH: 2087
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No: 235222.9
 US-60-324-185-16983

Alignment Scores:

Pred. No.: 2,096-173 Length: 2087
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-831-452-1 (1-424) x US-60-324-185-16983 (1-2087)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
 DB 189 ATGTCTGACCATGAGATGAGCTCCGCCGAAACCGGGTGAAGGCTCTCCAG 248
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 DB 249 CTGGGTAGTGGCGTAGAGGTGAATGAAGACATCCACCCGTCCGTCTCCCTCGA 308
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluAsnIleGluHisAlaPhe 60
 DB 309 GTTAAAGATTATCCGAATGCGATTCATCTCTAGAGAAAGCAATTGAACATGCTTC 368
 QY 61 IleuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 DB 369 ATCTCTATATACAGATGATATACGCTCTTATTAGAAATCAACAAACATCGAGATTAC 428
 QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100
 DB 429 AAATCTGCTGCATCTCCGAAAGAAAGACACAGATTAAGAAATTAAGAGATTGCAATT 488
 QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu 120
 DB 489 CCAAAACGAGAAAGCTCGAAGCGAGAGCTGTAAAGATATCCAAAGAAATTAACGAA 548
 QY 121 TyrAsnGluGluLysLysLysGluAlaGluLysLysAlaArgAsnMetAlaIleGlnGln 140
 DB 549 TATATGAGAGAAAGAAAGAGAGAGAGAGATTCGCCGGAACATGGCCATTCAGCAA 608
 QY 141 GluLeuGluLysGluLysGluArgValAlaGlnGlnLysGlnGlnGlnGluGlnGlu 160
 DB 609 GAGCTGAGAAAGAGAAAGAGAGAGAGAGAGATTCGCAATTTGGAAACAGAGAA 668
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgGluLys 180
 DB 669 CAGTTCATGCTTCGAGAGAGATGATCCGGAACAGAGATTAAGAAAGAGCACTGAAA 728
 QY 181 IleValGlnGluLysGluLysValAspProGluLysGluLysGluLysLysLysLys 200
 DB 729 ATTGACAGGAGATTGGAGAGTAGACCTGCTAGAGTGGCCCGGCTAGTGGCTGACTTG 768
 QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 DB 789 GAGAAAGCCCTCTTATGATGTGTTCCCACTTAAGCTCATTCATACAGCCCTTGAC 848
 QY 221 CysHisThrThrValArgProAlaLysProProValValAlaAspArgSerLeuLysProGly 240
 DB 849 TGTACACAACTGTAAGGCGAGCTAAGCACTGTGTGACAGAGCTCTTGAACCTCGA 908
 QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro 260
 DB 909 GCACTGAGCAACTCAGAAAGATATCCCAATCATGATGATTGCCCATGTGTGTGCT 968

QY 261 GlyArgLeuCysProGlnPheLeuGluLeuAlaSerAlaAsnThrValArgGlyValGlu 280
 DB 969 GGGCGGCTGTGCCACAGATTCTCCAGTTAGCAGTGGCAACACTGCCGGGGAGTGAG 1028
 QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 DB 1029 ACATGTGAAATTTCTGTGGAAGAACTGATGAGAAATTTACCATTAACCAATGTTTC 1068
 QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 DB 1089 ATCCCAAGCAAGAGCTGGGTCTGATTCAGCAACAGAGAAAGAAAGAACTTTTC 1148
 QY 321 IleuIleGlnAspGlnGluGluLysLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
 DB 1149 CTCATACAGAGATACAGAGGCTCATCACTGGGCTGATTCATACACCCACAG 1208
 QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro 360
 DB 1209 ACCGGTTTCTCTCCAGTCTGACCTTACACTGCTCTCTTACAGATGATGTGCA 1268
 QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380
 DB 1269 GAGTCAGTAGCCATGTGTGCTCCCAAGTCCAGAACTGAGATTTCTTAACACTACT 1328
 QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 DB 1329 GACCAATGACTAGAGAGAAATTTCTCTGTGCGAAGAAAGATTTATCCACAGCAG 1368
 QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
 DB 1389 GATCCACCTCTGTGTGTAGCTGACGACGAGTGTGTGACAGAGAGAGTGCATC 1448
 QY 421 ThrAspLeuArg 424
 DB 1449 ACAGACCTTCA 1460

RESULT 13

US-10-170-235-33766
 Sequence 33766, Application US/10170235
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig
 TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
 FILE REFERENCE: CL001380
 CURRENT APPLICATION NUMBER: US/10/170,235
 CURRENT FILING DATE: 2003-03-17
 NUMBER OF SEQ ID NOS: 42514
 SEQ ID NO 33766
 LENGTH: 2091
 TYPE: DNA
 ORGANISM: HUMAN
 US-10-170-235-33766

Alignment Scores:
 Pred. No.: 2,096-173 Length: 2091
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-831-452-1 (1-424) x US-10-170-235-33766 (1-2091)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
 DB 189 ATGTCTGACCATGAGATGAGCTCCGCCGAAACCGGGTGAAGGCTCTCCAG 248
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 DB 249 CTGGGTAGTGGCGTAGAGGTGAATGAAGACATTCACCCCGTCCGTACTTCCGCTCGA 308
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluAsnIleGluHisAlaPhe 60

```

Db      309 GTTGAGATTATCCGATGGCATCCATTACTGTGAGAAAGCAATGGAACATGCCCTTC 368
Qy      61 IleuTYranluTYrileThreuhellegluysleuproluysHiaArgAspTYr 80
Db      369 ATCCCTATACAGATATATACGCTCTTTATGTGAAACTACCAAAACATCGAGATTAC 428
Qy      81 LysSerAlaValIleProGluLysAspThrValLysLysLeuLysGluIleAlaPhe 100
Db      429 AAATCTGCTGCTATCTCGAAAAGAAAGACACAGTAAGAAATTAAGAGATTCGCTTT 488
Qy      101 ProlysaIaGluGluLeuLysAlaGluLeuLysArgTYrTYrLysGluTYrThrglu 120
Db      489 CCCAAAGCAGAGAGACTGAGGACAGAGCTGTAAACCATATACCAAGAAATATACAGAA 548
Qy      121 TYraEngluGluLysLysGluLysAlaGluGluLeuAlaArgAspMetAlaIleGln 140
Db      549 TATATGAAGAAAGAAAGAAAGCAAGCAAGATTTGCCCGAAACATGCGCATCCAGCA 608
Qy      141 GluLeuGluLysGluLysGluArgValAlaGlnGluLysGlnGlnLeuGluGln 160
Db      609 GAGCTGGAAGAGAAAACAGAGGCTAGCACACAGAGCAGCAGCAATTGAAACAGAA 668
Qy      161 GluPheHsaIaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluArgLeuLys 180
Db      669 CAGTCCCTGCTTGCAGAGATATGATCCGAAACAGAGAGCTAGAAAAGACGACTGAAA 728
Qy      181 IleValGlnGluPheGluLysValAspProGluLysGluLysGluProLeuValProAsp 200
Db      729 ATTGACAGAGATTGGAGAGTAGACCTTGCTAGTGGCCGCTAGTCCGACTTG 788
Qy      201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db      789 GAGAGAGCCCTCTTATGATGTGTCCCACTTAACAGTCTCATCAACAGCCTTCAGAC 848
Qy      221 CysHisThrThrValArgProAlaLysProProValAlaAspArgSerLeuLysProGly 240
Db      849 TGTACACACTGTAAAGCCACAGTACGCTGTGTGAGAGGCTCTTGAACCTGGA 908
Qy      241 AlaLeuSerAsnSerGluSerIleProThrIleAspGluLeuArgHsaValValPro 260
Db      909 GCACGTGAGCAACTCAGAAAGATATCCACATCATGATGCGCCATGTGCTGCT 968
Qy      261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerIleAsnThrAlaArgGlyValGlu 280
Db      969 GGGGCGCTGTCCACAGTTTCTCAGTTAGCCAGTCCACACTGCCGCGAGGTGAG 1028
Qy      281 ThrCysGluIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHsaValLeu 300
Db      1029 ACATGTGGAATTTCTGTGGAAGAACTGATGAGATATATTACATTACCACTTCCGCTTC 1088
Qy      301 IleProLysGlnSerAlaGlySerAspTYrCysAsnThrGluLysGluGluLeuPhe 320
Db      1089 ATCCCCAAGCAAAATGCTGGGTGATTACTGCAACACAGAGAAAGAAAGAACTTTTC 1148
Qy      321 LeuIleGlnAspGlnGlnGluLysLeuIleThrLeuGluTYrIleHsaThrHsaProThrGln 340
Db      1149 CTCTATACAGATACGAGGGCTCATACACTGGGCTGGATTCATATCCACCCCAACAG 1208
Qy      341 ThrIlePheLeuSerSerValAspLeuHisThrHisCysSerTYrGlnMetLeuPro 360
Db      1209 ACCCGCTTCTCTCCAGTGTGACCTACACATCAGCTCTTACCGAGATGTTGCCA 1268
Qy      361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGluPhePheLysLeuThr 380
Db      1269 GAGTCATACGCTATTTGCTCCCAAGTTCAGAGAACTGATTTCTTTAACTTACT 1328
Qy      381 AspHisGluLeuGluGluLysSerSerCysArgGluLysGluPheHisProHisSerLys 400
Db      1329 GACCATGAGCTAGAGGATTTCTCTGTCCAGAAAGATTTCTTCCACACAGAG 1388
Qy      401 AspProCysLeuPheCysSerCysSerHisValThrValAspArgAlaValIleThrIle 420

```

```

Db      1389 GATCCACCTCTGTTCTGTAGCTGACGACCACTGACTGTTGTGACAGAGCAGTGCACATC 1448
Qy      421 ThrAspLeuArg 424
Db      1449 ACAGACCTTCCA 1460

RESULT 14
US-09-606-776-4875
: Sequence 4875 Application US/09606776
: GENERAL INFORMATION:
: APPLICANT: Holtzman, Douglas A.
: APPLICANT: Myers, Paul
: APPLICANT: Gearing, David P.
: APPLICANT: Pan, Yacq
: TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
: FILE REFERENCE: 1600,1129-001
: CURRENT APPLICATION NUMBER: US/09/606,776
: PRIOR FILING DATE: 2000-06-27
: PRIOR APPLICATION NUMBER: 60/141,578
: PRIOR FILING DATE: 1999-06-29
: PRIOR APPLICATION NUMBER: 60/141,379
: PRIOR FILING DATE: 1999-06-28
: PRIOR APPLICATION NUMBER: 60/141,138
: PRIOR FILING DATE: 1999-06-28
: PRIOR APPLICATION NUMBER: 60/141,581
: NUMBER OF SEQ ID NOS: 5415
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 4875
: LENGTH: 2541
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-606-776-4875

Alignment Scores:
Pred. No.: 2,676-173 Length: 2541
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-831-452-1 (1-424) x US-09-606-776-4875 (1-2541)
Qy      1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
Db      621 ATGTCTGACCAACGAGATGTAGCTCCCGCCGAAAGCCGGGTAGGGCTCTCTCCAG 680
Qy      21 LeuGlySerAlaValAlaGluValAsnGluLysPheProProArgArgTYrPheArgSerGly 40
Db      681 CTGGGTAGTGGCGGTGAGGTGATGATGAGACATCCACCCCGTGGTACTTCGCTCTGGA 740
Qy      41 ValGluIleIleArgMetAlaSerIleTYrSerGluGluGluValAsnIleGluHisAlaPhe 60
Db      741 GTTGAGATTATCCGATGCGATCCATTTACTGTGAGAAAGCAATTCGAATCCCTTC 800
Qy      61 IleLeuTYranluTYrileThreuhellegluysleuproluysHiaArgAspTYr 80
Db      801 ATCCCTATACAGATATATACGCTCTTTATGAGAAACTACCAAAACATCGAGATTAC 860
Qy      81 LysSerAlaValIleProGluLysAspThrValLysLysLeuLysGluIleAlaPhe 100
Db      861 AAATCTGCTGCTATCTCGAAAAGAAAGCAAGTAAGAAATTAAGAGATTCGCTTT 920
Qy      101 ProlysaIaGluGluLeuLysAlaGluLeuLysArgTYrTYrLysGluTYrThrglu 120
Db      921 CCCAAAGCAGAGAGACTGAGGACAGAGCTGTAAACCATATACCAAGAAATATACAGAA 980
Qy      121 TYraEngluGluLysLysGluLysAlaGluGluLeuAlaArgAspMetAlaIleGln 140
Db      981 TATATGAAGAAAGAAAGAAAGCAAGCAAGATTTGCCCGAAACATGCGCATCCAGCAA 1040

```



```
Db      1401 GGGCGGCTGTGCCACAGATTCTCCAGTTAGCCAGTGCACAACACTGCCGGGGAGTGGAG 1460
QY      281  ThrCysglYlleuCYsgLYlyLeuMeArGnsngluphetHrIlethRhsValLeu 300
Db      1461 ACATGTGGAATTCTGTGTGGAATACTGATGGAATGATATTTACATTACCCATGTTCTC 1520
QY      301  IleProLYsgInSerAlaGLySerAspTYrCysAnthrGluAenGLuGLuLeuPhe 320
Db      1521 ATCCCAAGCAAAATGCTGGGTCTGATTACTGCACAACAGAGAACAGAAACTTTTC 1580
QY      321  LeuIleGlnAepGInGLyLeuIleThrLeuGLyTrpIleHsThrHsSProThrGln 340
Db      1581 CTCATTACAGAGATCAGCAGGGCCTCATCACACTGGGCTGGATTCTACTCACCCACACAG 1640
QY      341  ThrAlaPheLeuSerSerValAspLeuHsThrHsCysSerTYrGlnMetMetLeuPro 360
Db      1641 ACCCGTTTCTCTCCAGTGTGACCTACACACTCATCTCTTACAGATGATGTTGCCA 1700
QY      361  GluSerValAlaIleValCysSerProLYsPheGInGLuThrGLyPhePheLYsLeuThr 380
Db      1701 GAGTCAGTAGCCATTGTTTGCTCCCCCAAGTTCAGGAACCTGGAATCTTTAACTAAT 1760
QY      381  AspHsGLyLeuGLuGLuIleSerSerCysArgGLuLYsgLYPheHsSProHsSerLYs 400
Db      1761 GACCATGACTAGAGGAGATTCTTCCTGTGCCAGAAAGATTTCATCCACACAGCAAG 1820
QY      401  AspProPheLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
Db      1821 GATCCACTCTGTCTGTAGCTGACGCCACGTGACTGTGTGACAGACAGTGCATC 1880
QY      421  ThrAspLeuArg 424
Db      1881 ACAGACCTTCCA 1892
```

Search completed: August 23, 2004, 23:02:22
Job time : 4365 secs

This page Blank (uspio)

Tue Aug 24 10:03:55 2004

us-09-831-452-2.rnmp

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 13:02:40 ; Search time 6393 Seconds

(without alignments)
10512.355 Million cell updates/sec

Title: US-09-831-452-2
Sequence: 1 ctgtgcctcgtatgtctgacc.....ataagtaacaacttaactta 1910

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3757730 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/prodata/2/pna/PCRTUS_COMB.seq:*
2: /cgn2_6/prodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/prodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/prodata/2/pna/US08_COMB.seq:*
5: /cgn2_6/prodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/prodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/prodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/prodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/prodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/prodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/prodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/prodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/prodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/prodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/prodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/prodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/prodata/2/pna/US092B_COMB.seq:*
18: /cgn2_6/prodata/2/pna/US093A_COMB.seq:*
19: /cgn2_6/prodata/2/pna/US093B_COMB.seq:*
20: /cgn2_6/prodata/2/pna/US094_COMB.seq:*
21: /cgn2_6/prodata/2/pna/US095A_COMB.seq:*
22: /cgn2_6/prodata/2/pna/US095B_COMB.seq:*
23: /cgn2_6/prodata/2/pna/US095C_COMB.seq:*
24: /cgn2_6/prodata/2/pna/US095D_COMB.seq:*
25: /cgn2_6/prodata/2/pna/US096A_COMB.seq:*
26: /cgn2_6/prodata/2/pna/US096B_COMB.seq:*
27: /cgn2_6/prodata/2/pna/US096C_COMB.seq:*
28: /cgn2_6/prodata/2/pna/US096D_COMB.seq:*
29: /cgn2_6/prodata/2/pna/US096E_COMB.seq:*
30: /cgn2_6/prodata/2/pna/US097A_COMB.seq:*
31: /cgn2_6/prodata/2/pna/US097B_COMB.seq:*
32: /cgn2_6/prodata/2/pna/US097C_COMB.seq:*
33: /cgn2_6/prodata/2/pna/US098A_COMB.seq:*
34: /cgn2_6/prodata/2/pna/US098B_COMB.seq:*
35: /cgn2_6/prodata/2/pna/US098C_COMB.seq:*
36: /cgn2_6/prodata/2/pna/US098D_COMB.seq:*
37: /cgn2_6/prodata/2/pna/US099A_COMB.seq:*
38: /cgn2_6/prodata/2/pna/US099B_COMB.seq:*
39: /cgn2_6/prodata/2/pna/US099C_COMB.seq:*
40: /cgn2_6/prodata/2/pna/US099D_COMB.seq:*
41: /cgn2_6/prodata/2/pna/US099E_COMB.seq:*
42: /cgn2_6/prodata/2/pna/US099F_COMB.seq:*
43: /cgn2_6/prodata/2/pna/US100A_COMB.seq:*

44: /cgn2_6/prodata/2/pna/US100B_COMB.seq:*
45: /cgn2_6/prodata/2/pna/US101A_COMB.seq:*
46: /cgn2_6/prodata/2/pna/US101B_COMB.seq:*
47: /cgn2_6/prodata/2/pna/US102A_COMB.seq:*
48: /cgn2_6/prodata/2/pna/US102B_COMB.seq:*
49: /cgn2_6/prodata/2/pna/US103A_COMB.seq:*
50: /cgn2_6/prodata/2/pna/US103B_COMB.seq:*
51: /cgn2_6/prodata/2/pna/US104A_COMB.seq:*
52: /cgn2_6/prodata/2/pna/US104B_COMB.seq:*
53: /cgn2_6/prodata/2/pna/US106_COMB.seq:*
54: /cgn2_6/prodata/2/pna/US107A_COMB.seq:*
55: /cgn2_6/prodata/2/pna/US107B_COMB.seq:*
56: /cgn2_6/prodata/2/pna/US6000_COMB.seq:*
57: /cgn2_6/prodata/2/pna/US6001_COMB.seq:*
58: /cgn2_6/prodata/2/pna/US6002_COMB.seq:*
59: /cgn2_6/prodata/2/pna/US6003_COMB.seq:*
60: /cgn2_6/prodata/2/pna/US6004_COMB.seq:*
61: /cgn2_6/prodata/2/pna/US6005_COMB.seq:*
62: /cgn2_6/prodata/2/pna/US6006_COMB.seq:*
63: /cgn2_6/prodata/2/pna/US6007_COMB.seq:*
64: /cgn2_6/prodata/2/pna/US6008_COMB.seq:*
65: /cgn2_6/prodata/2/pna/US6009_COMB.seq:*
66: /cgn2_6/prodata/2/pna/US6010_COMB.seq:*
67: /cgn2_6/prodata/2/pna/US6011_COMB.seq:*
68: /cgn2_6/prodata/2/pna/US6012_COMB.seq:*
69: /cgn2_6/prodata/2/pna/US6013_COMB.seq:*
70: /cgn2_6/prodata/2/pna/US6014_COMB.seq:*
71: /cgn2_6/prodata/2/pna/US6015_COMB.seq:*
72: /cgn2_6/prodata/2/pna/US6016_COMB.seq:*
73: /cgn2_6/prodata/2/pna/US6017_COMB.seq:*
74: /cgn2_6/prodata/2/pna/US6018_COMB.seq:*
75: /cgn2_6/prodata/2/pna/US6019_COMB.seq:*
76: /cgn2_6/prodata/2/pna/US6020_COMB.seq:*
77: /cgn2_6/prodata/2/pna/US6021_COMB.seq:*
78: /cgn2_6/prodata/2/pna/US6022_COMB.seq:*
79: /cgn2_6/prodata/2/pna/US6023A_COMB.seq:*
80: /cgn2_6/prodata/2/pna/US6023B_COMB.seq:*
81: /cgn2_6/prodata/2/pna/US6024_COMB.seq:*
82: /cgn2_6/prodata/2/pna/US6025_COMB.seq:*
83: /cgn2_6/prodata/2/pna/US6026_COMB.seq:*
84: /cgn2_6/prodata/2/pna/US6027_COMB.seq:*
85: /cgn2_6/prodata/2/pna/US6028_COMB.seq:*
86: /cgn2_6/prodata/2/pna/US6029_COMB.seq:*
87: /cgn2_6/prodata/2/pna/US6030_COMB.seq:*
88: /cgn2_6/prodata/2/pna/US6031_COMB.seq:*
89: /cgn2_6/prodata/2/pna/US6032_COMB.seq:*
90: /cgn2_6/prodata/2/pna/US6033_COMB.seq:*
91: /cgn2_6/prodata/2/pna/US6034_COMB.seq:*
92: /cgn2_6/prodata/2/pna/US6035_COMB.seq:*
93: /cgn2_6/prodata/2/pna/US6036_COMB.seq:*
94: /cgn2_6/prodata/2/pna/US6037_COMB.seq:*
95: /cgn2_6/prodata/2/pna/US6038_COMB.seq:*
96: /cgn2_6/prodata/2/pna/US6039_COMB.seq:*
97: /cgn2_6/prodata/2/pna/US6040_COMB.seq:*
98: /cgn2_6/prodata/2/pna/US6041_COMB.seq:*
99: /cgn2_6/prodata/2/pna/US6042_COMB.seq:*
100: /cgn2_6/prodata/2/pna/US6043_COMB.seq:*
101: /cgn2_6/prodata/2/pna/US6044_COMB.seq:*
102: /cgn2_6/prodata/2/pna/US6045_COMB.seq:*
103: /cgn2_6/prodata/2/pna/US6046_COMB.seq:*
104: /cgn2_6/prodata/2/pna/US6047_COMB.seq:*
105: /cgn2_6/prodata/2/pna/US6048_COMB.seq:*
106: /cgn2_6/prodata/2/pna/US6049_COMB.seq:*
107: /cgn2_6/prodata/2/pna/US6050_COMB.seq:*
108: /cgn2_6/prodata/2/pna/US6051_COMB.seq:*
109: /cgn2_6/prodata/2/pna/US6052_COMB.seq:*
110: /cgn2_6/prodata/2/pna/US6053_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

QY 961 AGAATCTTCTCTCATACAGATGACAGAGGCTCATACATCTGGCTGATCTACTCA 1020
 Db 961 AGAATCTTCTCTCATACAGATGACAGAGGCTCATACATCTGGCTGATCTACTCA 1020
 QY 1021 CCCACACAGACCGCGTTCTCTCCAGTGTGACCTTACACATCTACGCTTACAGAT 1080
 Db 1021 CCCACACAGACCGCGTTCTCTCCAGTGTGACCTTACACATCTACGCTTACAGAT 1080
 QY 1081 GATGTGTGACAGAGTACAGTACATGTTGTCTCCCAAGTTCAGAGAACTGGATCTT 1140
 Db 1081 GATGTGTGACAGAGTACAGTACATGTTGTCTCCCAAGTTCAGAGAACTGGATCTT 1140
 QY 1141 TAAACTTAATGACATGACATGACAGAGATTTCTTCTGTGCGCAAGAAAGATTTCAATCC 1200
 Db 1141 TAAACTTAATGACATGACATGACAGAGATTTCTTCTGTGCGCAAGAAAGATTTCAATCC 1200
 QY 1201 ACACAGCAGAGATCCACTCTGTCTGTAGCTGACAGCAGTACTGTGTGACAGAGC 1260
 Db 1201 ACACAGCAGAGATCCACTCTGTCTGTAGCTGACAGCAGTACTGTGTGACAGAGC 1260
 QY 1261 AGTACCATCAACAGACCTTGATGAGCGTTGAGTCCAAACCTTCCAGAAACAAACAAA 1320
 Db 1261 AGTACCATCAACAGACCTTGATGAGCGTTGAGTCCAAACCTTCCAGAAACAAACAAA 1320
 QY 1321 CCATATCAGTGTACTGTAGCCCTTAATTAAGCTTTCTAGAAAGCTTTGGAGTTTGG 1380
 Db 1321 CCATATCAGTGTACTGTAGCCCTTAATTAAGCTTTCTAGAAAGCTTTGGAGTTTGG 1380
 QY 1381 TAGATAGTAAAGAGGAGGAGCTACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGA 1440
 Db 1381 TAGATAGTAAAGAGGAGGAGCTACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGA 1440
 QY 1441 AAGAAATTAATCTGAACATATTTTTTAGGCAAGTCAAGAAAGACATGCTCAACCCAAAGC 1500
 Db 1441 AAGAAATTAATCTGAACATATTTTTTAGGCAAGTCAAGAAAGACATGCTCAACCCAAAGC 1500
 QY 1501 AACTGTACTCAGAAATTAATGTAATCTCAGAAATTAATGAGTCTCAGAAATTAAGAAAGAT 1560
 Db 1501 AACTGTACTCAGAAATTAATGTAATCTCAGAAATTAATGAGTCTCAGAAATTAAGAAAGAT 1560
 QY 1561 GGTATTAATGAACCCCATATACCTTCTCTGTGATTCACCAATTTGTAACATTTTTTTC 1620
 Db 1561 GGTATTAATGAACCCCATATACCTTCTCTGTGATTCACCAATTTGTAACATTTTTTTC 1620
 QY 1621 CTCACAGTATCTCTTAATTTCTCTCAATTTGTAATTTGTTTATTTAATTTAATTTGAGCT 1680
 Db 1621 CTCACAGTATCTCTTAATTTCTCTCAATTTGTAATTTGTTTATTTAATTTAATTTGAGCT 1680
 QY 1681 CAATTAAGGCACTCTGTGAGAAATTTGGAAGCCATTTAGAAATCTTTTGGATTTTCTG 1740
 Db 1681 CAATTAAGGCACTCTGTGAGAAATTTGGAAGCCATTTAGAAATCTTTTGGATTTTCTG 1740
 QY 1741 TGGTTTATGGAATTAATGAAAGTCTTATTAATCTGGGTGAGGACACCTTACTTCAATTTG 1800
 Db 1741 TGGTTTATGGAATTAATGAAAGTCTTATTAATCTGGGTGAGGACACCTTACTTCAATTTG 1800
 QY 1801 ACCAGATTTGTTGGCTAAACATCCCGAAGATATTTTGTGAGAAATTAATTTGTAATTTA 1860
 Db 1801 ACCAGATTTGTTGGCTAAACATCCCGAAGATATTTTGTGAGAAATTAATTTGTAATTTA 1860
 QY 1861 ATAAATATTTGAGATATTTTCTCTCAATTAAGTAAAGTAAATTAATTTA 1910
 Db 1861 ATAAATATTTGAGATATTTTCTCTCAATTAAGTAAAGTAAATTAATTTA 1910

RESULT 2
 US-10-671-572A-2
 ; Sequence 2, Application US/10671572A
 ; GENERAL INFORMATION:
 ; APPLICANT: SUGAMURA, Kazuo
 ; APPLICANT: TANAKA, Nobuyuki
 ; TITLE OF INVENTION: Protein AMSH and cDNA thereof

; FILE REFERENCE: 2003-1382/MC/00653
 ; CURRENT APPLICATION NUMBER: US/10/671,572A
 ; CURRENT FILING DATE: 2003-09-29
 ; PRIOR APPLICATION NUMBER: 09/831,452
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: PCT/J99/06309
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: JP No. 10-322674
 ; PRIOR FILING DATE: 1998-11-12
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 2
 ; LENGTH: 1910
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 11..1282
 US-10-671-572A-2
 Query Match 100.0%; Score 1910; DB 53; Length 1910;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTTGGTCTGATGTGACCATGAGATGTGAGCTCCCGCCGAAAGACCGGGTGAAGGC 60
 Db 1 CTTGGTCTGATGTGACCATGAGATGTGAGCTCCCGCCGAAAGACCGGGTGAAGGC 60
 QY 61 TCTCTCCAGCTGGTGTAGTGCAGTGAAGTGAATGACATTTCAACCCCGTGGTACTT 120
 Db 61 TCTCTCCAGCTGGTGTAGTGCAGTGAAGTGAATGACATTTCAACCCCGTGGTACTT 120
 QY 121 CCGCTGTGAGATTTGATTTATCCGAATGCAATTTCTGTGAGAAAGGCAACATTGA 180
 Db 121 CCGCTGTGAGATTTGATTTATCCGAATGCAATTTCTGTGAGAAAGGCAACATTGA 180
 QY 181 ACAATGCTTATCTCTATTAACAAGTATCAAGCTCTTATTTAGAAATCTACCAACAA 240
 Db 181 ACAATGCTTATCTCTATTAACAAGTATCAAGCTCTTATTTAGAAATCTACCAACAA 240
 QY 241 TCGAGATTAACAATCTGCTGTCAATTCCTGAAGAAAGACACAGTAAAGAAATTAAGA 300
 Db 241 TCGAGATTAACAATCTGCTGTCAATTCCTGAAGAAAGACACAGTAAAGAAATTAAGA 300
 QY 301 GATTGCATTTCCCAAGACAGAAAGAGCTGAAGCAGAGCTGTAAACGATATACCAAGA 360
 Db 301 GATTGCATTTCCCAAGACAGAAAGAGCTGAAGCAGAGCTGTAAACGATATACCAAGA 360
 QY 361 ATATACAGATTAATGAAGAAAGAAAGAGACAGAGAAATTTGCCCGGAACATGGC 420
 Db 361 ATATACAGATTAATGAAGAAAGAAAGAGACAGAGAAATTTGCCCGGAACATGGC 420
 QY 421 CATTCAGCAAGAGCTGGAAAGAAAGAAAGAGAGGTGAACACAGAAACACAGAAAT 480
 Db 421 CATTCAGCAAGAGCTGGAAAGAAAGAAAGAGAGGTGAACACAGAAACACAGAAAT 480
 QY 481 GGAACAGAAAGATTCATGCTTGTGAGAGATGATCCGGAACCAAGAGCTGAGAAAGA 540
 Db 481 GGAACAGAAAGATTCATGCTTGTGAGAGATGATCCGGAACCAAGAGCTGAGAAAGA 540
 QY 541 GCGACTGAAATTTGTAAGAGAGTTTGGGAAGTGAACCTGAGCTTGAAGTGGCCGCTAGT 600
 Db 541 GCGACTGAAATTTGTAAGAGAGTTTGGGAAGTGAACCTGAGCTTGAAGTGGCCGCTAGT 600
 QY 601 GCCTGACTTGAAGAGCCCTCTTATGATGTGTCCCACTTAACAGTCAATCCATACA 660
 Db 601 GCCTGACTTGAAGAGCCCTCTTATGATGTGTCCCACTTAACAGTCAATCCATACA 660
 QY 661 GCCTTCAAGCTGTCAACAATGTAAGGCAAGTAAAGCCACTGAGTGAAGAGGCTCTT 720
 Db 661 GCCTTCAAGCTGTCAACAATGTAAGGCAAGTAAAGCCACTGAGTGAAGAGGCTCTT 720
 QY 721 GAAACCTGAGACCTGAGCACTCAGAAAGATTTCCCAATGATGATTCGCGCATGT 780

```

Db      721  GAAACCTGGAGACATGGAGCACTGAGAAATATCCACAAATGATGGGCGCATGT 780
Qy      781  GGTGTCCTGGGCGGCTGTGTCACAGTTTCTCCATTAGCCAGTSCCAACCTGCGC 840
Db      781  GGTGTCCTGGGCGGCTGTGTCACAGTTTCTCCATTAGCCAGTSCCAACCTGCGC 840
Qy      841  GGGAGTGGAGACATGTGGAAATCTGTGTGAAAACTGATGAGAAATTAATTCATTAC 900
Db      841  GGGAGTGGAGACATGTGGAAATCTGTGTGAAAACTGATGAGAAATTAATTCATTAC 900
Qy      901  CCATGTCCTCATCCCAAGCAAAATGCTGGGTCTGATTATTCGCAACAGAGACGAGA 960
Db      901  CCATGTCCTCATCCCAAGCAAAATGCTGGGTCTGATTATTCGCAACAGAGACGAGA 960
Qy      961  AGAATCTTTCTCTACAGATGAGATGAGAGGCGCTCATCAGCTGGGCTGATTAATCTCA 1020
Db      961  AGAATCTTTCTCTACAGATGAGATGAGAGGCGCTCATCAGCTGGGCTGATTAATCTCA 1020
Qy      1021  CCCCACACAGACGCGGTTTCTCTCAGTGTGACCTTCAACATCAGCTGCTTAACAGAT 1080
Db      1021  CCCCACACAGACGCGGTTTCTCTCAGTGTGACCTTCAACATCAGCTGCTTCAACAGAT 1080
Qy      1081  GATGTCGACAGTCAAGTCAAGGCAATGTTGTCTCCCAAGTCCAGGAACTGATTTCTT 1140
Db      1081  GATGTCGACAGTCAAGTCAAGGCAATGTTGTCTCCCAAGTCCAGGAACTGATTTCTT 1140
Qy      1141  TAAATCTATGACATGATGATGAGAGATTTCTTCTGTCTGCGCAAGAAAGATTTGATTC 1200
Db      1141  TAAATCTATGACATGATGATGAGAGATTTCTTCTGTCTGCGCAAGAAAGATTTGATTC 1200
Qy      1201  ACACAGACAGGATCCACTCTGTCTGATCTGACGACGATGATCTTTGTGACAGAGC 1260
Db      1201  ACACAGACAGGATCCACTCTGTCTGATCTGACGACGATGATCTTTGTGACAGAGC 1260
Qy      1261  AGTACCATCAACAGACCTTGATGAGCGTTTGAATCCAAACCTTCCAAAGAACAAAA 1320
Db      1261  AGTACCATCAACAGACCTTGATGAGCGTTTGAATCCAAACCTTCCAAAGAACAAAA 1320
Qy      1321  CCATATCAGTGTATCTGAGCGGCTTAATTAATTAAGCTTCTGAAAGCTTTGAAATTTTG 1380
Db      1321  CCATATCAGTGTATCTGAGCGGCTTAATTAATTAAGCTTCTGAAAGCTTTGAAATTTTG 1380
Qy      1381  TAGATAGTAAAGGGGCGGATCAGCTGAGAAAGCTGATTTGTATTTCAAGTTTGA 1440
Db      1381  TAGATAGTAAAGGGGCGGATCAGCTGAGAAAGCTGATTTGTATTTCAAGTTTGA 1440
Qy      1441  AAGAAATTAATCAATATTTTAAAGCAAGTCAAGAAAGCAATGTCACCCAAAGC 1500
Db      1441  AAGAAATTAATCAATATTTTAAAGCAAGTCAAGAAAGCAATGTCACCCAAAGC 1500
Qy      1501  AACTGTAACTCAGAAATTAAGTTACTCAGAAATTAAGTACTCAGAAATTAAGAAAT 1560
Db      1501  AACTGTAACTCAGAAATTAAGTTACTCAGAAATTAAGTACTCAGAAATTAAGAAAT 1560
Qy      1561  GGTATTAATGAACCCCAATATACCTTCTTCTGATTCACCAATGTATAATTTTTC 1620
Db      1561  GGTATTAATGAACCCCAATATACCTTCTTCTGATTCACCAATGTATAATTTTTC 1620
Qy      1621  CTCACAGTATCTCTTAATTTCTCTTAATTTCAATTTGTTATATTTTCTCTGGGCT 1680
Db      1621  CTCACAGTATCTCTTAATTTCTCTTAATTTCAATTTGTTATATTTTCTCTGGGCT 1680
Qy      1681  CAATTAAGGGGATCTGTGCAAAATTTGAAAGCCATTTAGAAATCTTTTGATTTTCCTG 1740
Db      1681  CAATTAAGGGGATCTGTGCAAAATTTGAAAGCCATTTAGAAATCTTTTGATTTTCCTG 1740
Qy      1741  TGGTTAATGGAATATGAATGAGCTTATTAAGGGGTGAGGAGCACTTACTCCATTGG 1800
Db      1741  TGGTTAATGGAATATGAATGAGCTTATTAAGGGGTGAGGAGCACTTACTCCATTGG 1800
Qy      1801  ACCGATTTGTTGGCTAACATCCGAGAAATGATTTTGTCAAGAAATTAATTTATTTA 1860

```

```

Db      1801  ACCGATTTGTTGGCTAACATCCGAGAAATGATTTTGTCAAGAAATTAATTTATTTA 1860
Qy      1861  ATTAATTAATTAAGATATTTTCTTCAATTAATTAAGATTAATTAATTAATTAATTA 1910
Db      1861  ATTAATTAATTAAGATATTTTCTTCAATTAATTAAGATTAATTAATTAATTAATTA 1910

RESULT 3
US-09-770-175-8814
; Sequence 8814, Application US/09770175
; GENERAL INFORMATION:
; APPLICANT: Geary, David P.
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2058-001
; CURRENT APPLICATION NUMBER: US/09/770.175
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,874
; NUMBER OF SEQ ID NOS: 8967
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8814
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-770-175-8814

Query Match      100.0%; Score 1910; DB 32; Length 2001;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  CTGTGCTCTGATGATGTGACATGAGATGTAGGCTCCCGCCGAAAGCCGGGTGAGGCG 60
Db      56  CTGTGCTCTGATGATGTGACATGAGATGTAGGCTCCCGCCGAAAGCCGGGTGAGGCG 115
Qy      61  TCTCTCCAGCTGGTATGCGGTGAGGTGAATGAAGACATTCACCCCGTCTACTT 120
Db      116  TCTCTCCAGCTGGTATGCGGTGAGGTGAATGAAGACATTCACCCCGTCTACTT 175
Qy      121  CCGCTCTGAGTGAATTAATCCGATGATCCATTTACTCTGAGAAAGCAACATTTGA 180
Db      176  CCGCTCTGAGTGAATTAATCCGATGATCCATTTACTCTGAGAAAGCAACATTTGA 235
Qy      181  ACATGCTTCAATCTCTTAATTAACAGTATATACGCTCTTATTAAGAAATTAACCA 240
Db      236  ACATGCTTCAATCTCTTAATTAACAGTATATACGCTCTTATTAAGAAATTAACCA 295
Qy      241  TCGAATTAACAAATCTGCTCATTCCTGAAAGAAAGACACAGTAAAGAAATTAAGGA 300
Db      296  TCGAATTAACAAATCTGCTCATTCCTGAAAGAAAGACACAGTAAAGAAATTAAGGA 355
Qy      301  GATTGCAATTTCCCAAGCAGAGCTGAGGCAAGCTGTTAAAGCATATACCAAGA 360
Db      356  GATTGCAATTTCCCAAGCAGAGCTGAGGCAAGCTGTTAAAGCATATACCAAGA 415
Qy      361  ATATACAGATTAATTAAGAAAGAAAGAAAGAAAGAAATTTGGCCCGGAAATATGCG 420
Db      416  ATATACAGATTAATTAAGAAAGAAAGAAAGAAAGAAATTTGGCCCGGAAATATGCG 475
Qy      421  CATCAGCAGAGCTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 480
Db      476  CATCAGCAGAGCTGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 535
Qy      481  GGACAGGAAAGATTCATGCTCTTCTGAGAGATATTCGGAACAGAGAGCTTAAGAAA 540
Db      536  GGACAGGAAAGATTCATGCTCTTCTGAGAGATATTCGGAACAGAGAGCTTAAGAAA 595
Qy      541  GCGACTGAAATTTGATGAGAGTTTGGGAAAGCTGAGACCTGAGTGGCCCGCTAGT 600
Db      596  GCGACTGAAATTTGATGAGAGTTTGGGAAAGCTGAGACCTGAGTGGCCCGCTAGT 655
Qy      601  GCGTACTTGGAGAGCCCTCTTATGATGTGTTCCCACTTAACAGCTTCATCATACA 660

```

Db 656 GCGTCACTGGAGAGAGCCCTCCTTAGATGTGTTCCCAACCTTAACAATCTCATCCATACACA 715
Qy 661 GCCTTCAGACTGTACACAATCTAAGGCCAGCTAAGCAACCTGTGGAGCAGGTCTT 720
Db 716 GCTTCAGACTGTACACAATCTAAGGCCAGCTAAGCAACCTGTGTGGAGCAGGTCTT 775
Qy 721 GAAACCTGGAGCACTGAGCAACTCAGAAAGTATTCACAAATCGATGATGGCCATGT 780
Db 776 GAAACCTGGAGCACTGAGCAACTCAGAAAGTATTCACAAATCGATGATGGCCATGT 835
Qy 781 GGTGGTGGCTGGGGGGGGCTGTGGCCACAGTTCTCCAGTTAGCCGAGGCCAAGCTGGCCG 840
Db 836 GGTGGTGGCTGGGGGGGGCTGTGGCCACAGTTCTCCAGTTAGCCGAGGCCAAGCTGGCCG 895
Qy 841 GGGAGTGGAGACATGTGGAATTTCTGTGAGAAACCTGATGAGATGAATTTACATTAAC 900
Db 896 GGGAGTGGAGACATGTGGAATTTCTGTGAGAAACCTGATGAGATGAATTTACATTAAC 955
Qy 901 CCATGTTCTCATCCCAAGCAAGAGTGGTGGTCTGATTTACTGCAACACAGAAACGAAGA 960
Db 956 CCATGTTCTCATCCCAAGCAAGAGTGGTGGTCTGATTTACTGCAACACAGAAACGAAGA 1015
Qy 961 AGAATCTTCTCTCATPACAGATGACAGAGGAGCTCATGACACTGGGGCTGGATTCACTCA 1020
Db 1016 AGAATCTTCTCTCATPACAGATGACAGAGGAGCTCATGACACTGGGGCTGGATTCACTCA 1075
Qy 1021 CCCCACACAGACCCGCTTCTCTCAGTGTGACCTTACACACTGACTCTTAACAGAT 1080
Db 1076 CCCCACACAGACCCGCTTCTCTCAGTGTGACCTTACACACTGACTCTTAACAGAT 1135
Qy 1081 GATGTGGCAGAGTCAAGTACAGATTTGTTGCTCCCCCAAGTTCAGAGAACTGATTTCTT 1140
Db 1136 GATGTGGCAGAGTCAAGTACAGATTTGTTGCTCCCCCAAGTTCAGAGAACTGATTTCTT 1195
Qy 1141 TAAACTAATGACCATGAGCTAGAGAGATTTCTTCTGTGCGCAGAAAGGATTTCAATCC 1200
Db 1196 TAAACTAATGACCATGAGCTAGAGAGATTTCTTCTGTGCGCAGAAAGGATTTCAATCC 1255
Qy 1201 ACAACAGCAAGATTCACCTCTGTTCTGTAGCTGACGCCACTGACTGTTGTGACAGAGC 1260
Db 1256 ACAACAGCAAGATTCACCTCTGTTCTGTAGCTGACGCCACTGACTGTTGTGACAGAGC 1315
Qy 1261 AGTGACCATTCACAGACCTTGATGAGAGGTTGATGCCAACCTCCAGAGAACCAAAA 1320
Db 1316 AGTGACCATTCACAGACCTTGATGAGAGGTTGATGCCAACCTCCAGAGAACCAAAA 1375
Qy 1321 CCATATCAGTGTACTGTAGCCCTTAATTTAAGCTTTCTAGAAAAGCTTTGAGAACTTTTG 1380
Db 1376 CCATATCAGTGTACTGTAGCCCTTAATTTAAGCTTTCTAGAAAAGCTTTGAGAACTTTTG 1435
Qy 1381 TAGATAGTAAAGAGGGGGGAGTCACTGAGAAAGAGCTGATTTGTATTTTCAAGTTTGA 1440
Db 1436 TAGATAGTAAAGAGGGGGGAGTCACTGAGAAAGAGCTGATTTGTATTTTCAAGTTTGA 1495
Qy 1441 AAGAAATTAATGAAATTTTAAAGCAAGTCAAGAAAGAAAGCATGCTACCCAAAAGC 1500
Db 1496 AAGAAATTAATGAAATTTTAAAGCAAGTCAAGAAAGAAAGCATGCTACCCAAAAGC 1555
Qy 1501 AACTGTAACTAGAAATTAAGTACTCAGAAATTAAGTACTGAAATTAAGAAAGAT 1560
Db 1556 AACTGTAACTAGAAATTAAGTACTCAGAAATTAAGTACTGAAATTAAGAAAGAT 1615
Qy 1561 GGTATTAATGAACCCCATATACCTTCTGTGATTCACCAATTTGTAACATTTTTC 1620
Db 1616 GGTATTAATGAACCCCATATACCTTCTGTGATTCACCAATTTGTAACATTTTTC 1675
Qy 1621 CTCTCAGTATCTCTTAATTTCTCTTAATTTGATTTGTTAATTTTACCTTGGGCT 1680
Db 1676 CTCTCAGTATCTCTTAATTTCTCTTAATTTGATTTGTTAATTTTACCTTGGGCT 1735
Qy 1681 CAATAGGCAATCTGTGAGAAATTTGAGCAATTTTGAAGAACTTTTGGATTTTCTG 1740

Db 1736 CAATAGGCAATCTGTGAGAAATTTGAGCAATTTTGAAGCAATTTTGAAGAACTTTTCTG 1795
Qy 1741 TGGTTATGCAATATGAATGAGCTTATTTACTGTGGGTGAGAGCAAGCTTACTCCATTG 1800
Db 1796 TGGTTATGCAATATGAATGAGCTTATTTACTGTGGGTGAGAGCAAGCTTACTCCATTG 1855
Qy 1801 ACCAGTTGTTGGCTTAACACATCCGAAAGATGATTTTGTCAAGATTAATTTATTTA 1860
Db 1856 ACCAGTTGTTGGCTTAACACATCCGAAAGATGATTTTGTCAAGATTAATTTATTTA 1915
Qy 1861 ATTAATTTTCAAGATTAATTTTCTCTACATTAATTAATTTTCAATTTAATTTA 1910
Db 1916 ATTAATTTTCAAGATTAATTTTCTCTACATTAATTAATTTAATTTAATTTA 1965

RESULT 4
US-10-752-442-1
; Sequence 1: Application US/10752442
; GENERAL INFORMATION:
; APPLICANT: Itoh, Fumiko
; APPLICANT: Itoh, Susumu
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: Itoh, Peter
; TITLE OF INVENTION: SWAP ASSOCIATING POLYPEPTIDES
; FILE REFERENCE: L0461.70096US01
; CURRENT APPLICATION NUMBER: US/10/752,442
; CURRENT FILING DATE: 2004-01-06
; PRIOR APPLICATION NUMBER: US 09/665,479
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US 60/154,846
; PRIOR FILING DATE: 1999-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)..(1334)
US-10-752-442-1

Query Match 100.0%; Score 1910; DB 55; Length 2003;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTGTCCTGATGTCTGACCAATGAGATGTGAGCTCCCGCCGGAAGACCGGGTGAGGCG 60
Db 53 CTTGTCCTGATGTCTGACCAATGAGATGTGAGCTCCCGCCGGAAGACCGGGTGAGGCG 112
Qy 61 TCTCTCCAGCTGGGTAGTGGGTAGAGGTGATGAAGCATTTCAACCCCGTCGGTACTT 120
Db 113 TCTCTCCAGCTGGGTAGTGGGTAGAGGTGATGAAGCATTTCAACCCCGTCGGTACTT 172
Qy 121 CCGCTGTGAGTGAATTTCCGAATGCGATCAATTTACTGTGAGAGAGCAACATTGA 180
Db 173 CCGCTGTGAGTGAATTTCCGAATGCGATCAATTTACTGTGAGAGAGCAACATTGA 232
Qy 181 ACAATGCTTCACTCTCTAATGAAGATATCAAGCTCTTATTTGAGAACTTCCAAAACA 240
Db 233 ACAATGCTTCACTCTCTAATGAAGATATCAAGCTCTTATTTGAGAACTTCCAAAACA 292
Qy 241 TCGAGATTACAAATCTGTCTGATTTCTGTAAGAAAGAAAGCAAGTAAAGAAATTAAGA 300
Db 293 TCGAGATTACAAATCTGTCTGATTTCTGTAAGAAAGAAAGCAAGTAAAGAAATTAAGA 352
Qy 301 GATTGATTTCCCAAGCAAGAGAGCTGAAGCAAGAGCTGTTAAACGATTAACCAAGAA 360
Db 353 GATTGATTTCCCAAGCAAGAGAGCTGAAGCAAGAGCTGTTAAACGATTAACCAAGAA 412
Qy 361 ATATACAGAAATTAATGAAGAAAGAAAGAGAGAGAGATTTGGCCCGAACAATGCG 420
Db 413 ATATACAGAAATTAATGAAGAAAGAAAGAGAGAGAGATTTGGCCCGAACAATGCG 472

QY 421 CATCAGCAGAGCTGGAAAAAGAAAAAGAGGAGTACCAACAGAGAGCAGCAATT 480
DB 473 CATCAGCAGAGAGCTGGAAAAAGAAAAAGAGGAGTACCAACAGAGAGCAGCAATT 532
QY 481 GGAACAGGAACAGTTCATGCTTCGAGAGATGATCGGAAACAGAGAGCTAGAAAAAGA 540
DB 533 GGAACAGGAACAGTTCATGCTTCGAGAGATGATCGGAAACAGAGAGCTAGAAAAAGA 592
QY 541 GCGACTGAGAAATTTGACAGAGATTGGGAAAGTAAAGCCCTGAGCTAGTGGCCCTAGT 600
DB 593 GCGACTGAGAAATTTGACAGAGATTGGGAAAGTAAAGCCCTGAGCTAGTGGCCCTAGT 652
QY 601 GCGACTGAGAAAGCCCTGAGTAAAGTTCCTCCCACTTAAACATCTCATCCATAC 660
DB 653 GCGACTGAGAAAGCCCTGAGTAAAGTTCCTCCCACTTAAACATCTCATCCATAC 712
QY 661 GCGTTCAGACTGTGACAACTGTAAAGCCAGCTAAAGCCAGCTGTGTGAGAGTCTT 720
DB 713 GCGTTCAGACTGTGACAACTGTAAAGCCAGCTAAAGCCAGCTGTGTGAGAGTCTT 772
QY 721 GAAACCTGAGAGCATGAGCAACTAGAAAGTATTCACCAATGATGATGGCCATGT 780
DB 773 GAAACCTGAGAGCATGAGCAACTAGAAAGTATTCACCAATGATGATGGCCATGT 832
QY 781 GGTGTGCTGGGCGGCTGTGCTCCAGCTTCTCCAGTTAGCCAGTSCCAACCTGCGC 840
DB 833 GGTGTGCTGGGCGGCTGTGCTCCAGCTTCTCCAGTTAGCCAGTSCCAACCTGCGC 892
QY 841 GGGAGTGGAGACATGTGAAATTTCTGTGAAAACTGATGAGATGATTTACCAATTAC 900
DB 893 GGGAGTGGAGACATGTGAAATTTCTGTGAAAACTGATGAGATGATTTACCAATTAC 952
QY 901 CCAGTTCCTATCCCAAGCAAGTGTGGGTGATTAATGCAACAGAGAGAGAG 960
DB 953 CCAGTTCCTATCCCAAGCAAGTGTGGGTGATTAATGCAACAGAGAGAGAG 1012
QY 961 AGAATTTCTCTATAGAGATGAGAGAGGCTTCATCAGCTGGCTGATTAATCTCA 1020
DB 1013 AGAATTTCTCTATAGAGATGAGAGAGGCTTCATCAGCTGGCTGATTAATCTCA 1072
QY 1021 CCCACACAGAGCCGCTTCTCTCCAGTGTGACCTTCAACATCACTGCTCTTAACAGAT 1080
DB 1073 CCCACACAGAGCCGCTTCTCTCCAGTGTGACCTTCAACATCACTGCTCTTAACAGAT 1132
QY 1081 GATGTGCGAGAGTCAAGTATGATTTGCTCCCAAGTTCAGAGAACTGGATTCTT 1140
DB 1133 GATGTGCGAGAGTCAAGTATGATTTGCTCCCAAGTTCAGAGAACTGGATTCTT 1192
QY 1141 TAAACTACTGACCATGAGATGAGAGATTTCTTCTGTGCGCAAGAAAGATTATTC 1200
DB 1193 TAAACTACTGACCATGAGATGAGAGATTTCTTCTGTGCGCAAGAAAGATTATTC 1252
QY 1201 ACACAGAGAGATCCACTCTGTCTGTAGCTGAGCAGTGAATGTTGTGACAGAGC 1260
DB 1253 ACACAGAGAGATCCACTCTGTCTGTAGCTGAGCAGTGAATGTTGTGACAGAGC 1312
QY 1261 AGTACCATGACAGACCTTGATGAGCGTTGATCCAAACACTTCCAAAGAACAAAA 1320
DB 1313 AGTACCATGACAGACCTTGATGAGCGTTGATCCAAACACTTCCAAAGAACAAAA 1372
QY 1321 CCAATATAGTGTATGAGCCCTTAATTAAGCTTCTGAGAAAGCTTGGAGTTTG 1380
DB 1373 CCAATATAGTGTATGAGCCCTTAATTAAGCTTCTGAGAAAGCTTGGAGTTTG 1432
QY 1381 TAGATAGTAAAGGGGGGATCACTGAGAAAGCTGATTTTGTATTTCAAGTTTGA 1440
DB 1433 TAGATAGTAAAGGGGGGATCACTGAGAAAGCTGATTTTGTATTTCAAGTTTGA 1492
QY 1441 AAGAAATTAATCAATATTTTAAAGCAAGTCAAGAAAGAAACATGTGTACCCAAAGC 1500
DB 1493 AAGAAATTAATCAATATTTTAAAGCAAGTCAAGAAAGAAACATGTGTACCCAAAGC 1552

QY 1501 AACTGTAATCTAGAAATTAAGTTACTAGAAATTAAGTACTAGAAATTAAGAAAGAT 1560
DB 1553 AACTGTAATCTAGAAATTAAGTTACTAGAAATTAAGTACTAGAAATTAAGAAAGAT 1612
QY 1561 GGTATATGAACCCCATATACCTTCTCTGATGATGACCAATGTTAAACATTTTTC 1620
DB 1613 GGTATATGAACCCCATATACCTTCTCTGATGATGACCAATGTTAAACATTTTTC 1672
QY 1621 CTCTGAGTATCTCTTAATTTCTCTTAATTTCAATTTGTTATTTTACCCTGGGCT 1680
DB 1673 CTCTGAGTATCTCTTAATTTCTCTTAATTTCAATTTGTTATTTTACCCTGGGCT 1732
QY 1681 CAATAAGGCACTGTGAGAAATTTGAGAGCATTGAGAAATCTTTGATTTTCTG 1740
DB 1733 CAATAAGGCACTGTGAGAAATTTGAGAGCATTGAGAAATCTTTGATTTTCTG 1792
QY 1741 TGGTTATGGCAATATGATGAGCTTATTAATGAGGAGAGCAAGTTACTCAATTG 1800
DB 1793 TGGTTATGGCAATATGATGAGCTTATTAATGAGGAGAGCAAGTTACTCAATTG 1852
QY 1801 ACCAGATTTGGGTACATCCCGAAGATGATTTTGCAGAAATTAATTTATTTA 1860
DB 1853 ACCAGATTTGGGTACATCCCGAAGATGATTTTGCAGAAATTAATTTATTTA 1912
QY 1861 ATAAATATTTGAGATATTTTCTCTCAATTAAGTAACATTAATTA 1910
DB 1913 ATAAATATTTGAGATATTTTCTCTCAATTAAGTAACATTAATTA 1962

RESULT 5
US-09-881-797-3931/c
; Sequence 3931, Application US/09881797
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: H0001-04
; CURRENT APPLICATION NUMBER: US/09/881,797
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/211,749
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3931
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Human
US-09-881-797-3931

Query Match 100.0%; Score 1910; DB 36; Length 2045;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 1939 CTGGTCTGATGCTGACATGAGATGTAGGCTCCCGCCGAGAGCGGTGAGGCG 60
QY 61 TCTCTCCAGCTGGGTAGTGGGTGAGATGATGAGATTCACCCCTGGTACTT 120
DB 1879 TCTCTCCAGCTGGGTAGTGGGTGAGATGATGAGATTCACCCCTGGTACTT 1820
QY 121 CCGCTCTGAGTGAATTAATCCGAATGAGATTCATTAATCTGAGAGAGGCAATGTA 180
DB 1819 CCGCTCTGAGTGAATTAATCCGAATGAGATTCATTAATCTGAGAGAGGCAATGTA 1760
QY 181 ACATGCTTCAATCTCTATTAACAGATATACAGCTCTTTATGAGAACTACCAAAACA 240
DB 1759 ACATGCTTCAATCTCTATTAACAGATATACAGCTCTTTATGAGAACTACCAAAACA 1700
QY 241 TCGAGATTACAAATCTGTGATCTCTGAGAAAGAAAGACAGATTAAGAAATTTAAAGA 300
DB 1699 TCGAGATTACAAATCTGTGATCTCTGAGAAAGAAAGACAGATTAAGAAATTTAAAGA 1640

QY 301 GATTGATTTCCCAAGCAGAGAGTGAAGGAGAGCTGTAAACGATATACCAAGA 360
Db 1639 GATTGATTTCCCAAGCAGAGAGTGAAGGAGAGCTGTAAACGATATACCAAGA 1580
QY 361 ATATACGATATATATGAAGAAAAGAAAGAAAGCAAGAAATGGCCCGGAACATGGC 420
Db 1579 ATATACGATATATATGAAGAAAAGAAAGAAAGCAAGAAATGGCCCGGAACATGGC 1520
QY 421 CATCCACCAAGAGCTGGAAAAAGAAAACAGAGGAGTCAACAACAGAAAGCAGCAATT 480
Db 1519 CATCCACCAAGAGCTGGAAAAAGAAAACAGAGGAGTCAACAACAGAAAGCAGCAATT 1460
QY 481 GAAACAGAAACAGTTCATGCTTCGAGAGATGATCCGAAACAGAGAGCTAGAAAAAGA 540
Db 1459 GAAACAGAAACAGTTCATGCTTCGAGAGATGATCCGAAACAGAGAGCTAGAAAAAGA 1400
QY 541 GCGACTGAAATTTGTACAGAGAGTTGGGAAGTGAAGCCCTGGCTAGTGGCCCGCTAGT 600
Db 1399 GCGACTGAAATTTGTACAGAGAGTTGGGAAGTGAAGCCCTGGCTAGTGGCCCGCTAGT 1340
QY 601 GCGTGAATTTGGAGAGCCCTCTTGTAGATGTGTTCCCACTTAACAGTCTCATCCATCA 660
Db 1339 GCGTGAATTTGGAGAGCCCTCTTGTAGATGTGTTCCCACTTAACAGTCTCATCCATCA 1280
QY 661 GCGTTCAGACTGTACACAACTGTGAAGCCAGCTAAGCCACTGTGTGAGCAGGTCTT 720
Db 1279 GCGTTCAGACTGTACACAACTGTGAAGCCAGCTAAGCCACTGTGTGAGCAGGTCTT 1220
QY 721 GAAACCTGGAGCACTGGACAACTCAGAAAGATTTCCACATCGATGGATTTGGCCCTGT 780
Db 1219 GAAACCTGGAGCACTGGACAACTCAGAAAGATTTCCACATCGATGGATTTGGCCCTGT 1160
QY 781 GGTGTGCTCTGGAGGCTGTGTGCCACAGTTTCTCCAGTTAGCCAGTGCACACTGCCG 840
Db 1159 GGTGTGCTCTGGAGGCTGTGTGCCACAGTTTCTCCAGTTAGCCAGTGCACACTGCCG 1100
QY 841 GGGAGTGGAGACATGTGGAAATTTCTGTGGAAAACTGATAGAGATTAATTTCCATTAC 900
Db 1099 GGGAGTGGAGACATGTGGAAATTTCTGTGGAAAACTGATAGAGATTAATTTCCATTAC 1040
QY 901 CCATGTTCTATCCCAAGCAAGAGTGTGGGTCTGATTAAGTCAACAGAGAAAGCAAGA 960
Db 1039 CCATGTTCTATCCCAAGCAAGAGTGTGGGTCTGATTAAGTCAACAGAGAAAGCAAGA 980
QY 961 AGAACTTTTCTCTATACAGATACGAGGCTCATACACTGGGCTGATTCATCTCA 1020
Db 979 AGAACTTTTCTCTATACAGATACGAGGCTCATACACTGGGCTGATTCATCTCA 920
QY 1021 CCCACACAGAGCCGCTTCTCTCCAGTGTGACCTCAACACTGCTCTTACAGAT 1080
Db 919 CCCACACAGAGCCGCTTCTCTCCAGTGTGACCTCAACACTGCTCTTACAGAT 860
QY 1081 GATGTGCCAGAGTCACTAGCCATTTGTTGCTCCCAAGTTCAGAGAACTGATTTCTT 1140
Db 859 GATGTGCCAGAGTCACTAGCCATTTGTTGCTCCCAAGTTCAGAGAACTGATTTCTT 800
QY 1141 TAAACTATACGACATGAGCTAGAGAGATTTCTTCTGTGCGCAGAGAAAGATTTCAATC 1200
Db 799 TAAACTATACGACATGAGCTAGAGAGATTTCTTCTGTGCGCAGAGAAAGATTTCAATC 740
QY 1201 ACAAGCAGAGATCACTCTGTTCTGTAGCTGACAGCAGTGAATGTTGACAGAGC 1260
Db 739 ACAAGCAGAGATCACTCTGTTCTGTAGCTGACAGCAGTGAATGTTGACAGAGC 680
QY 1261 AGTACCATACAGACCTTGTGAGAGGCTTGAATCCAAACCTTCCAAAGAAACAAAGAA 1320
Db 679 AGTACCATACAGACCTTGTGAGAGGCTTGAATCCAAACCTTCCAAAGAAACAAAGAA 620
QY 1321 CCATATCAGTGTAGTACCCCTTAATTTAAGCTTTCTAAGAAAGCTTTGGAAGTTTGG 1380
Db 619 CCATATCAGTGTAGTACCCCTTAATTTAAGCTTTCTAAGAAAGCTTTGGAAGTTTGG 560
QY 1381 TAGATAGTAGAAGGGGGGAGCTACCTGAGAAAGAGCTGATTTTGTATTTCAAGTTTGA 1440

Db 559 TAGATAGTAGAAGGGGGGAGCTACCTGAGAAAGAGCTGATTTGTATTTAGTTTGA 500
QY 1441 AAGAAATTAATCAATATTTTGTAGCAAGTCAAGAAAGAAACATGGCAGCCAAAAGC 1500
Db 499 AAGAAATTAATCAATATTTTGTAGCAAGTCAAGAAAGAAACATGGCAGCCAAAAGC 440
QY 1501 AACTGTACTCAGAAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGAAAGAT 1560
Db 439 AACTGTACTCAGAAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGAAAGAT 380
QY 1561 GGTATATGACACCCCATATACCTTCTCTGATTCACCAATTTGTTAATTTTTC 1620
Db 379 GGTATATGACACCCCATATACCTTCTCTGATTCACCAATTTGTTAATTTTTC 320
QY 1621 CTCTGACTATCTCTATATTTCTCTAATTTCAATTTGTTAATTTAATTTAATTTT 1680
Db 319 CTCTGACTATCTCTATATTTCTCTAATTTCAATTTGTTAATTTAATTTAATTTT 260
QY 1681 CATTAAGGCACTGTGACAGAAATTTGGAAGCCATTTAGAAATCTTTGATTTCTG 1740
Db 259 CATTAAGGCACTGTGACAGAAATTTGGAAGCCATTTAGAAATCTTTGATTTCTG 200
QY 1741 TGGTTATAGCAATATGAATGAGCTTATTAATGGGGGAGAGGAGAGCTTCAATTTG 1800
Db 199 TGGTTATAGCAATATGAATGAGCTTATTAATGGGGGAGAGGAGAGCTTCAATTTG 140
QY 1801 ACCAGATTTGTTGGTAAACACATCCGGAAGATGATTTGTGAGAAATTAATTTA 1860
Db 139 ACCAGATTTGTTGGTAAACATCCGGAAGATGATTTGTGAGAAATTAATTTA 80
QY 1861 ATTAATATTTCAATATTTTCTCTACATTAAGTAATTAATTTA 1910
Db 79 ATTAATATTTCAATATTTTCTCTACATTAAGTAATTAATTTA 30

RESULT 6
US-10-021-698-3931/c
; Sequence 3931, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: H0001-04
; CURRENT APPLICATION NUMBER: US/10/021,698
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3931
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Human
US-10-021-698-3931

Query Match 100.0%; Score 1910; DB 43; Length 2045;
Best Local Similarity 100.0%; Pct. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGGTCTGATGTGACCATGAGATGTGAGCTCCCGCCGAAGACCGGGTGAAGGC 60
Db 1939 CTTGGTCTGATGTGACCATGAGATGTGAGCTCCCGCCGAAGACCGGGTGAAGGC 1880
QY 61 TCTTCCCAAGCTGGGTATGCGGTGAGATGAATGAATTAACATTTCCACCCCGCTGACTT 120
Db 1879 TCTTCCCAAGCTGGGTATGCGGTGAGATGAATGAATTAACATTTCCACCCCGCTGACTT 1820
QY 121 CCGCTCTGAGTGTGATTAATCCGAATGCAATTCATTTACTGTGAGAAAGCAATTTGA 180
Db 1819 CCGCTCTGAGTGTGATTAATCCGAATGCAATTCATTTACTGTGAGAAAGCAATTTGA 1760
QY 181 ACATGCTTCAATCTCTATACCAAGTATATACGCTCTTTATGAAACTACCAAAACA 240


```

Db      1759  ACATGCTTATCTCTATATACAGATATATACGCTTTTATGAGAACTACCAAAACA 1700
Qy      241   TCGAGATTACAAAATCTGCTGTCATTTCTGAAAAAGAACACAGTAAGAAATTAAGA 300
Db      1699  TCGAGATTACAAAATCTGCTGTCATTTCTGAAAAAGAACACAGTAAGAAATTAAGA 1640
Qy      301   GATTGCTTTCCCAAGACAGAGACTGAGAGGCTTTAAACGATATCCAAAGA 360
Db      1639  GATTGCTTTCCCAAGACAGAGACTGAGAGGCTTTAAACGATATCCAAAGA 1580
Qy      361   ATATACGAAATATATATAGAAAAAGAAAGAGAGAGAAATTTGGCCCGGAACATGGC 420
Db      1579  ATATACGAAATATATATAGAAAAAGAAAGAGAGAGAAATTTGGCCCGGAACATGGC 1520
Qy      421   CATTCACAGAAAGCTGGAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db      1519  CATTCACAGAAAGCTGGAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1460
Qy      481   GGAACAGAGAAAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      1459  GGAACAGAGAAAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1400
Qy      541   GCGACTGAAAAATTTGATACAGAGAGTTTGGAGAGGTAGACCTTGAGTGGCCGCTAGT 600
Db      1399  GCGACTGAAAAATTTGATACAGAGAGTTTGGAGAGGTAGACCTTGAGTGGCCGCTAGT 1340
Qy      601   GCGTGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      1339  GCGTGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
Qy      661   GCGTTGACAGCTGTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db      1279  GCGTTGACAGCTGTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
Qy      721   GAAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db      1219  GAAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1160
Qy      781   GGTGTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db      1159  GGTGTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1100
Qy      841   GGGAGTGAAGACATGTGAAATTTCTGTGAAAACTGATGAGAGAAATTTACACTTAC 900
Db      1099  GGGAGTGAAGACATGTGAAATTTCTGTGAAAACTGATGAGAGAAATTTACACTTAC 1040
Qy      901   CCAGTCTCATATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db      1039  CCAGTCTCATATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 980
Qy      961   AGAAGCTTTCTCATACAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db      979   AGAAGCTTTCTCATACAGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 920
Qy      1021  CCCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db      919   CCCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 860
Qy      1081  GATGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db      859   GATGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 800
Qy      1141  TAAACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db      799   TAAACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740
Qy      1201  AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db      739   AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 680
Qy      1261  AGTACCATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

```

```

Db      679   AGTACCATCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 620
Qy      1321  CCATATCAGTGTACTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db      619   CCATATCAGTGTACTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 560
Qy      1381  TAGATAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db      559   TAGATAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500
Qy      1441  AAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db      499   AAGAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 440
Qy      1501  AACTGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db      439   AACTGTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380
Qy      1561  GGTATATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db      379   GGTATATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 320
Qy      1621  CTCTCAGTATCTCTTAATTTCTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 1680
Db      319   CTCTCAGTATCTCTTAATTTCTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 260
Qy      1681  CAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db      259   CAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 200
Qy      1741  TGGTTATGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db      199   TGGTTATGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 140
Qy      1801  ACCGATTTGTTGGCTAACATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db      139   ACCGATTTGTTGGCTAACATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 80
Qy      1861  ATAAATATTTAGAGATATTTTCTCTTACAAATTAAGTAACATTAACCTTA 1910
Db      79   ATAAATATTTAGAGATATTTTCTCTTACAAATTAAGTAACATTAACCTTA 30

RESULT 7
US-10-170-235-33766
; Sequence 33766, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: CLO01380
; CURRENT APPLICATION NUMBER: US/10/170,235
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 33766
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-33766

Query Match      100.0%; Score 1910; DB 46; Length 2091;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Db 299 CCGCTCGAGATTGAGATTCAGAAATGGCATTCATTTACTGTGAGAAAGCAACATTGA 358
 Qy 181 ACATGCTCTCATCTCTATATACAGATATATACAGCTCTTTATTTAGAAATCTACCAAAACA 240
 Db 359 ACATGCTCTCATCTCTATATACAGATATATACAGCTCTTTATTTAGAAATCTACCAAAACA 418
 Qy 241 TCGAGATTACAAATCTGTCTGTCATTTCTGAAAAGAAAGACAGTAAGAAATTTAAAGA 300
 Db 419 TCGAGATTACAAATCTGTCTGTCATTTCTGAAAAGAAAGACAGTAAGAAATTTAAAGA 478
 Qy 301 GATTGCAATTTCCCAAGCAGAAAGCTGAGGCGAGCTGTTAAACGATATACCAAGA 360
 Db 479 GATTGCAATTTCCCAAGCAGAAAGCTGAGGCGAGCTGTTAAACGATATACCAAGA 538
 Qy 361 ATATACAGAAATATATATAGTAAAGAAAGAAAGACAGAGGAATTTGCGCGGAACATGGC 420
 Db 539 ATATACAGAAATATATATAGTAAAGAAAGAAAGACAGAGGAATTTGCGCGGAACATGGC 598
 Qy 421 CATCCACGAAAGCTGGAAGAAAGAAAGACAGAGGATGCAACAAGAGCAGACGCAATT 480
 Db 599 CATCCACGAAAGCTGGAAGAAAGAAAGACAGAGGATGCAACAAGAGCAGACGCAATT 658
 Qy 481 GGACAGAAACAGTTTCATGCTTCGAGAGATGATTCGGAACGAGAGCTAGAAAAAGA 540
 Db 659 GGACAGAAACAGTTTCATGCTTCGAGAGATGATTCGGAACGAGAGCTAGAAAAAGA 718
 Qy 541 GCGACTGAAAATTTGATAGAGAGTTGGGAAAGTGAACCTTGAGCTAGAGTGGCGGCTAGT 600
 Db 719 GCGACTGAAAATTTGATAGAGAGTTGGGAAAGTGAACCTTGAGCTAGAGTGGCGGCTAGT 778
 Qy 601 GCGTCACTTGGAGAAAGCCCTCTTGAATGTGTTCCTCCACCTTAAACGATCTCATCATACA 660
 Db 779 GCGTCACTTGGAGAAAGCCCTCTTGAATGTGTTCCTCCACCTTAAACGATCTCATCATACA 838
 Qy 661 GCGTTCAGACTGTGACACAACTGTAAAGCCAGCTAAACCACTGTGTGTGAGCAAGTCTT 720
 Db 839 GCGTTCAGACTGTGACACAACTGTAAAGCCAGCTAAACCACTGTGTGTGAGCAAGTCTT 898
 Qy 721 GAAACCTGAGAGCTGAGCAACTGAGAAAGTATTCACCAATGATGATGGGCCATGT 780
 Db 899 GAAACCTGAGAGCTGAGCAACTGAGAAAGTATTCACCAATGATGATGGGCCATGT 958
 Qy 781 GGTGTGCTGAGGAGGCTGTGAGCCACAGTTTCTCCAGTTTGGAGAGCCCAACCTGCGCG 840
 Db 959 GGTGTGCTGAGGAGGCTGTGAGCCACAGTTTCTCCAGTTTGGAGAGCCCAACCTGCGCG 1018
 Qy 841 GGGAGTGGAGACATGTGAAATTTCTGTGAGAAACTGATGAGATGAATTTACCATTAAC 900
 Db 1019 GGGAGTGGAGACATGTGAAATTTCTGTGAGAAACTGATGAGATGAATTTACCATTAAC 1078
 Qy 901 CCATGTTCTCATCTCCCAAGCAAAAGTGTGGGTCTGATTAATGCAACACAGAAACGAAAGA 960
 Db 1079 CCATGTTCTCATCTCCCAAGCAAAAGTGTGGGTCTGATTAATGCAACACAGAAACGAAAGA 1138
 Qy 961 AGAATTTTCTCATACAGATCAGAGGCTCATACACTGGGCTGGAGTTTATACTCA 1020
 Db 1139 AGAATTTTCTCATACAGATCAGAGGCTCATACACTGGGCTGGAGTTTATACTCA 1198
 Qy 1021 CCCCACACAGACCGCGTTTCTCTCAAGTGTGACCTACACACTCACTGCTCTTACAGAT 1080
 Db 1199 CCCCACACAGACCGCGTTTCTCTCAAGTGTGACCTACACACTCACTGCTCTTACAGAT 1258
 Qy 1081 GATGTTCCAGAGTCACTAGCAATGTTTGTCTCCCAAGTCTTCCAGAGAACTGGATTCTT 1140
 Db 1259 GATGTTCCAGAGTCACTAGCAATGTTTGTCTCCCAAGTCTTCCAGAGAACTGGATTCTT 1318
 Qy 1141 TAAACTTAAGTCACTAGCAATGAGAGATTTCTTCTGTGCGCAGAAAGATTTTATCC 1200
 Db 1319 TAAACTTAAGTCACTAGCAATGAGAGATTTCTTCTGTGCGCAGAAAGATTTTATCC 1378
 Qy 1201 ACAAGAGAGATCACTGCTGTTCTGTAGCTCAGACAGTACTGTTGTGACAGAGC 1260

Db 1379 ACACAGCAAGAGATCCACTCTGTTCTGTAGCTGACAGCCAGTGACTGTTGGACAGAGC 1438
 Qy 1261 AGTACCATCACAGACCTTCGATGAGGCTTTAGATCCCAACACTTCCAGAAACAAACAAA 1320
 Db 1439 AGTACCATCACAGACCTTCGATGAGGCTTTAGATCCCAACACTTCCAGAAACAAACAAA 1498
 Qy 1321 CCATATCACTGATCTGATGAGGCTTTAATTAAGCTTTTCAAGAGCTTTGAAAGTTTGG 1380
 Db 1499 CCATATCACTGATCTGATGAGGCTTTAATTAAGCTTTTCAAGAGCTTTGAAAGTTTGG 1558
 Qy 1381 TGAATGTGAAAAGGGGCGCATACCTGAGAAAGAGCTGATTTTGTATTCAGCTTTGAA 1440
 Db 1559 TGAATGTGAAAAGGGGCGCATACCTGAGAAAGAGCTGATTTTGTATTCAGCTTTGAA 1618
 Qy 1441 AAGAAATTAAGTGAACATATTTTGTAGGCAAGTACAGAAAGAAACATGTCACCCAAAGC 1500
 Db 1619 AAGAAATTAAGTGAACATATTTTGTAGGCAAGTACAGAAAGAAACATGTCACCCAAAGC 1678
 Qy 1501 AACTGTAAGTGAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGTAAAGAAAGAT 1560
 Db 1679 AACTGTAAGTGAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGTAAAGAAAGAT 1738
 Qy 1561 GGTATTAAGTAAAGGCGCATACCTCTCTGATGATGACCAATGTTTAACTTTTTC 1620
 Db 1739 GGTATTAAGTAAAGGCGCATACCTCTCTGATGATGACCAATGTTTAACTTTTTC 1798
 Qy 1621 CTCTGAGTATCTCTTATATTTCTCTATATTTCAATTTGTTATTTACTCTGAGGCT 1680
 Db 1799 CTCTGAGTATCTCTTATATTTCTCTATATTTCAATTTGTTATTTACTCTGAGGCT 1858
 Qy 1681 CAATAAGGCGCATCTGTGCAAAATTTGGAAGCAATTTAGAAATCTTTGATTTTCTG 1740
 Db 1859 CAATAAGGCGCATCTGTGCAAAATTTGGAAGCAATTTAGAAATCTTTGATTTTCTG 1918
 Qy 1741 TGGTTATGCAATTTGATGATGAGCTTATTAATGAGGAGAGAGCTTACTCCATTTG 1800
 Db 1919 TGGTTATGCAATTTGATGATGAGCTTATTAATGAGGAGAGAGCTTACTCCATTTG 1978
 Qy 1801 ACCAGATGTTTGGCTTAACACATCCCGAAGAAAGATTTTGTGAGAAATTAATTTATTA 1860
 Db 1979 ACCAGATGTTTGGCTTAACACATCCCGAAGAAAGATTTTGTGAGAAATTAATTTATTA 2038
 Qy 1861 ATAAATATTTGAGATTTTCTCTCAATAAAGTAAATTAATTA 1910
 Db 2039 ATAAATATTTGAGATTTTCTCTCAATAAAGTAAATTAATTA 2088

RESULT 8
 US-09-606-776-4875
 ; Sequence 4875, Application US/09606776
 ; GENERAL INFORMATION:
 ; APPLICANT: Holzman, Douglas A.
 ; APPLICANT: Myers, Paul
 ; APPLICANT: Gearing, David P.
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; TITLE OF INVENTION: THEREFOR
 ; FILE REFERENCE: 1600.1129-001
 ; CURRENT APPLICATION NUMBER: US/09/606,776
 ; PRIOR FILING DATE: 2000-06-27
 ; PRIOR APPLICATION NUMBER: 60/141,578
 ; PRIOR FILING DATE: 1999-06-29
 ; PRIOR APPLICATION NUMBER: 60/141,379
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: 60/141,138
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: 60/141,581
 ; PRIOR FILING DATE: 1999-06-29
 ; NUMBER OF SEQ ID NOS: 5415
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4875
 ; LENGTH: 2541
 ; TYPE: DNA

ORGANISM: Homo sapiens

US-09-606-776-4875

Query Match 100.0%; Score 1910; DB 25; Length 2541;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTTGGTCTGATGTCTGACCATGAGATGTGAGCTTCCGCCCGAAGACCGGGTGAAGGC 60
DB CTTGGTCTGATGTCTGACCATGAGATGTGAGCTTCCGCCCGAAGACCGGGTGAAGGC 670
QY 61 TCTCTCCAGCTGGGTAGTCCGGTAGAGTGAATGAAGACATTCACCCCGTGGTACTT 120
DB TCTCTCCAGCTGGGTAGTCCGGTAGAGTGAATGAAGACATTCACCCCGTGGTACTT 120
QY 671 TCTCTCCAGCTGGGTAGTCCGGTAGAGTGAATGAAGACATTCACCCCGTGGTACTT 730
DB TCTCTCCAGCTGGGTAGTCCGGTAGAGTGAATGAAGACATTCACCCCGTGGTACTT 730
QY 121 CCGCTCTGAGTGAATTAATCCGATGGCATTCATTTACTCTGAGAAAGGCAACATTTGA 180
DB CCGCTCTGAGTGAATTAATCCGATGGCATTCATTTACTCTGAGAAAGGCAACATTTGA 180
QY 731 CCGCTCTGAGTGAATTAATCCGATGGCATTCATTTACTCTGAGAAAGGCAACATTTGA 790
DB CCGCTCTGAGTGAATTAATCCGATGGCATTCATTTACTCTGAGAAAGGCAACATTTGA 790
QY 181 ACATGCTTCACTGCTATTAACAAGTATATCAAGCTCTTATTTGAGAAACTATCAAAAGA 240
DB ACATGCTTCACTGCTATTAACAAGTATATCAAGCTCTTATTTGAGAAACTATCAAAAGA 240
QY 791 ACATGCTTCACTGCTATTAACAAGTATATCAAGCTCTTATTTGAGAAACTATCAAAAGA 850
DB ACATGCTTCACTGCTATTAACAAGTATATCAAGCTCTTATTTGAGAAACTATCAAAAGA 850
QY 241 TCGAGATTACAAATCTGCTGTCTATTCCTGAAAAAGAACACAGTAAGATTTAAAGGA 300
DB TCGAGATTACAAATCTGCTGTCTATTCCTGAAAAAGAACACAGTAAGATTTAAAGGA 300
QY 851 TCGAGATTACAAATCTGCTGTCTATTCCTGAAAAAGAACACAGTAAGATTTAAAGGA 910
DB TCGAGATTACAAATCTGCTGTCTATTCCTGAAAAAGAACACAGTAAGATTTAAAGGA 910
QY 301 GATTGCTTTCCCAAGACAAAGAGCTGAGAGAGCTGTTAAAGCATATACCAAGA 360
DB GATTGCTTTCCCAAGACAAAGAGCTGAGAGAGCTGTTAAAGCATATACCAAGA 360
QY 911 GATTGCTTTCCCAAGACAAAGAGCTGAGAGAGCTGTTAAAGCATATACCAAGA 970
DB GATTGCTTTCCCAAGACAAAGAGCTGAGAGAGCTGTTAAAGCATATACCAAGA 970
QY 361 ATATACAGAAATATATATGAGAAAAAGAAAGAAAGAGAGAGATTTGGCCGGAACATGAC 420
DB ATATACAGAAATATATATGAGAAAAAGAAAGAAAGAGAGAGATTTGGCCGGAACATGAC 420
QY 971 ATATACAGAAATATATATGAGAAAAAGAAAGAAAGAGAGATTTGGCCGGAACATGAC 1030
DB ATATACAGAAATATATATGAGAAAAAGAAAGAAAGAGAGATTTGGCCGGAACATGAC 1030
QY 421 CATTCACAGAAAGCTGAGAAAGAGAAAGAGAGAGTATGCAACACAGAGAGAGAGATTT 480
DB CATTCACAGAAAGCTGAGAAAGAGAAAGAGAGAGTATGCAACACAGAGAGAGAGATTT 480
QY 1031 CATTCACAGAAAGCTGAGAAAGAGAAAGAGAGAGTATGCAACACAGAGAGAGATTT 1090
DB CATTCACAGAAAGCTGAGAAAGAGAAAGAGAGAGTATGCAACACAGAGAGAGATTT 1090
QY 481 GGAACAGAGAAAGCTGATGCTCTCGAGAGATGATCCGGAACAGAGAGATTTAGAAAAAGA 540
DB GGAACAGAGAAAGCTGATGCTCTCGAGAGATGATCCGGAACAGAGAGATTTAGAAAAAGA 540
QY 1091 GGAACAGAGAAAGCTGATGCTCTCGAGAGATGATCCGGAACAGAGAGATTTAGAAAAAGA 1150
DB GGAACAGAGAAAGCTGATGCTCTCGAGAGATGATCCGGAACAGAGAGATTTAGAAAAAGA 1150
QY 541 GCGACTGAAAAATTTGACAGAGATTTGGGAAAGTATGACCTGCGCTGAGTGGCCGCTAGT 1210
DB GCGACTGAAAAATTTGACAGAGATTTGGGAAAGTATGACCTGCGCTGAGTGGCCGCTAGT 1210
QY 601 GCGCTGAGTGGAGAGGCTCTCTGATGATGTTCCCACTTAACAGTCTCATCCATACA 660
DB GCGCTGAGTGGAGAGGCTCTCTGATGATGTTCCCACTTAACAGTCTCATCCATACA 660
QY 1211 GCGTGACTTGGAGAGGCTCTCTGATGATGTTCCCACTTAACAGTCTCATCCATACA 1270
DB GCGTGACTTGGAGAGGCTCTCTGATGATGTTCCCACTTAACAGTCTCATCCATACA 1270
QY 661 GCGTTCAGACTGTTCACAACTGTATGAGCCAGCTTAAGCCACTGTGTGAGCAGGTCTT 720
DB GCGTTCAGACTGTTCACAACTGTATGAGCCAGCTTAAGCCACTGTGTGAGCAGGTCTT 720
QY 721 GAAACCTTGAGACATGAGCAATCTGAGAAAGTATTTCCCAATGATGATTTGGCCATGT 780
DB GAAACCTTGAGACATGAGCAATCTGAGAAAGTATTTCCCAATGATGATTTGGCCATGT 780
QY 1331 GAAACCTTGAGACATGAGCAATCTGAGAAAGTATTTCCCAATGATGATTTGGCCATGT 1390
DB GAAACCTTGAGACATGAGCAATCTGAGAAAGTATTTCCCAATGATGATTTGGCCATGT 1390
QY 781 GGTGTGCTGCTGGGCGGCTGTGCCCCAGATTTCCTCAAGTATGAGCCCAACATGCGCG 840
DB GGTGTGCTGCTGGGCGGCTGTGCCCCAGATTTCCTCAAGTATGAGCCCAACATGCGCG 840
QY 1391 GGTGTGCTGCTGGGCGGCTGTGCCCCAGATTTCCTCAAGTATGAGCCCAACATGCGCG 1450
DB GGTGTGCTGCTGGGCGGCTGTGCCCCAGATTTCCTCAAGTATGAGCCCAACATGCGCG 1450
QY 841 GGGAGTGGAGACATGTGGAAATCTCTGTGAAAACTGATGAGAAATGATTTCCATTCAC 900
DB GGGAGTGGAGACATGTGGAAATCTCTGTGAAAACTGATGAGAAATGATTTCCATTCAC 900
QY 1451 GGGAGTGGAGACATGTGGAAATCTCTGTGAAAACTGATGAGAAATGATTTCCATTCAC 1510
DB GGGAGTGGAGACATGTGGAAATCTCTGTGAAAACTGATGAGAAATGATTTCCATTCAC 1510
QY 901 CCAATGTTCTATCCCAAGCAAGTGTGGGTGATTTATTCGACACACAGAGACAGAGA 960
DB CCAATGTTCTATCCCAAGCAAGTGTGGGTGATTTATTCGACACACAGAGACAGAGA 960
QY 1511 CCAATGTTCTATCCCAAGCAAGTGTGGGTGATTTATTCGACACACAGAGACAGAGA 1570
DB CCAATGTTCTATCCCAAGCAAGTGTGGGTGATTTATTCGACACACAGAGACAGAGA 1570
QY 961 AGAATTTTCTCTCAATACAGATGACAGAGGCTCATACACTGGGCTGATTTACTACTCA 1020
DB AGAATTTTCTCTCAATACAGATGACAGAGGCTCATACACTGGGCTGATTTACTACTCA 1020
```

```
DB 1571 AGAATTTTCTCTCAATACAGATGACAGAGGCTCATACACTGGGCTGATTTACTACTCA 1630
QY 1021 CCCACACACACCGCGTTTCTCTCCAGTGTGAGCTTACACATTCAGTCTTACAGAT 1080
DB 1631 CCCACACACACCGCGTTTCTCTCCAGTGTGAGCTTACACATTCAGTCTTACAGAT 1690
QY 1081 GATGTGCGAGAGTCAATGAGCATTTGTTGCTCCCGCAAGTCTCCAGAACTGATTTCT 1140
DB 1681 GATGTGCGAGAGTCAATGAGCATTTGTTGCTCCCGCAAGTCTCCAGAACTGATTTCT 1750
QY 1141 TAACTTAATGACATGAGCATGAGAGATTTCTTCTGTGTGCGCAAGAAAGATTTGATCC 1200
DB 1751 TAACTTAATGACATGAGCATGAGAGATTTCTTCTGTGTGCGCAAGAAAGATTTGATCC 1810
QY 1201 AACAGCAAGATCAACCTGTGTGTGAGCTGAGAGCAAGTACTGTTGTGACAGAGC 1260
DB 1811 AACAGCAAGATCAACCTGTGTGTGAGCTGAGAGCAAGTACTGTTGTGACAGAGC 1870
QY 1261 AGTACCATCAACAGCTTGTGATGAGCGTTGAGTCCAACTTCCAAAGCAACAAA 1320
DB 1871 AGTACCATCAACAGCTTGTGATGAGCGTTGAGTCCAACTTCCAAAGCAACAAA 1930
QY 1321 CCAATGAGTGTACTGTAGCCCTTAATTTAGCTTTCTAAGAAAGCTTTGGAAGTTTG 1380
DB 1931 CCAATGAGTGTACTGTAGCCCTTAATTTAGCTTTCTAAGAAAGCTTTGGAAGTTTG 1990
QY 1381 TAGATAGTAAAGAGGAGGATCACTGAGAAAGCTGATTTTGTATTTCAAGTTTGA 1440
DB 1991 TAGATAGTAAAGAGGAGGATCACTGAGAAAGCTGATTTTGTATTTCAAGTTTGA 2050
QY 1441 AAGAAATTAATGAACTATTTTATGAGCAAGTCAAGAAAGAAAGAGTCAACCAAAAGC 1500
DB 2051 AAGAAATTAATGAACTATTTTATGAGCAAGTCAAGAAAGAAAGAGTCAACCAAAAGC 2110
QY 2051 AAGAAATTAATGAACTATTTTATGAGCAAGTCAAGAAAGAAAGAGTCAACCAAAAGC 2110
DB 2111 AAGTATTAATGAACTATTTTATGAGCAAGTCAAGAAAGAAAGAGTCAACCAAAAGC 2170
QY 1561 GGTATATGAACCCCATATACCTTCTCTGATTCAGCAATTTGTTAATTTTCTTTC 1620
DB 2171 GGTATATGAACCCCATATACCTTCTCTGATTCAGCAATTTGTTAATTTTCTTTC 2230
QY 1621 GGTATATGAACCCCATATACCTTCTCTGATTCAGCAATTTGTTAATTTTCTTTC 1680
DB 2231 GGTATATGAACCCCATATACCTTCTCTGATTCAGCAATTTGTTAATTTTCTTTC 2290
QY 1681 CAATGAGGATCTGTGCAAGAAATTTGGAAGCAATTTGAAATCTTTTGAATTTCTG 1740
DB 2291 CAATGAGGATCTGTGCAAGAAATTTGGAAGCAATTTGAAATCTTTTGAATTTCTG 2350
QY 1741 TGGTTATGGAATATGAGAGCTTATTAATGAGGAGAGAGAGAGTCTCATTTG 1800
DB 2351 TGGTTATGGAATATGAGAGCTTATTAATGAGGAGAGAGAGTCTCATTTG 2410
QY 2411 ACCAGATTTGTTGGTAAACATCCCGAAGATGATTTGTGAGAAATTAATTTGATTA 1860
DB 2471 ACCAGATTTGTTGGTAAACATCCCGAAGATGATTTGTGAGAAATTAATTTGATTA 2470
QY 1861 ATAAATTAATTTGAGATATTTTCTCTTACATTAAGTAAATTAATTAATTA 1910
DB 2471 ATAAATTAATTTGAGATATTTTCTCTTACATTAAGTAAATTAATTAATTA 2520
```

RESULT 9
US-09-652-917-3220
Sequence 3220, Application US/09652917
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
APPLICANT: Holzman, Douglas A.
APPLICANT: Disefano, Peter
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600, 1170-001

/ CURRENT APPLICATION NUMBER: US/09/652,917
/ CURRENT FILING DATE: 2000-08-30
/ PRIOR APPLICATION NUMBER: 60/151,422
/ PRIOR FILING DATE: 1999-08-30
/ NUMBER OF SEQ ID NOS: 3855
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 3220
/ LENGTH: 2541
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-652-917-3220

Query Match 100.0%; Score 1910; DB 27; Length 2541;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGCTCTGATGTGTGACATGAGATGTGAGCTCCGCCGGAAGACCGGGTGAAGGC 60
DB 611 CTTGCTCTGATGTGTGACATGAGATGTGAGCTCCGCCGGAAGACCGGGTGAAGGC 670
QY 61 TCTCTCCAGCTGGGTAGTGGCGTGAAGTGAATGAAGACATTCACCCCGTGGTACTT 120
DB 671 TCTCTCCAGCTGGGTAGTGGCGTGAAGTGAATGAAGACATTCACCCCGTGGTACTT 730
QY 121 CCGCTCTGAGTTGAGATTATCCGAATGCGATTCATTACTCTGAGGAAGGCAATTGA 180
DB 731 CCGCTCTGAGTTGAGATTATCCGAATGCGATTCATTACTCTGAGGAAGGCAATTGA 790
QY 181 ACATGCGCTCATCCTCTATTAACAAGTATATACGCTCTTTATTAGAACTACCAACA 240
DB 791 ACATGCGCTCATCCTCTATTAACAAGTATATACGCTCTTTATTAGAACTACCAACA 850
QY 241 TCGAGATTACAATCTGTCTCATTCCTGAAAGAAAGACACAGTAAAGAAATTAAAGA 300
DB 851 TCGAGATTACAATCTGTCTCATTCCTGAAAGAAAGACACAGTAAAGAAATTAAAGA 910
QY 301 GATTGCTATTTCCCAAGCAGAGAGCTGAAGGCGAGCTGTTAAACGATATACCAAGA 360
DB 911 GATTGCTATTTCCCAAGCAGAGAGCTGAAGGCGAGCTGTTAAACGATATACCAAGA 970
QY 361 ATATACAGATATATATGAGAAAGAAAGAGAGAGAGAGTGGCCCGGAACATAGC 420
DB 971 ATATACAGATATATATGAGAAAGAAAGAGAGAGAGTGGCCCGGAACATAGC 1030
QY 421 CATCCAGCAAGAGCTGGAAGAAAGAAACAGAGGGTACACACAGAGAGAGCAATT 480
DB 1031 CATCCAGCAAGAGCTGGAAGAAAGAAACAGAGGGTACACACAGAGAGAGCAATT 1090
QY 481 GGAACAGGAACAGTTTCATGCTTCGAGAGATGATCCGGAACAGAGAGCTAAGAAAGA 540
DB 1091 GGAACAGGAACAGTTTCATGCTTCGAGAGATGATCCGGAACAGAGAGCTAAGAAAGA 1150
QY 541 GCGACTGAAATTTGTACAGAGGTTTGGGAAGTAGACCTGGCTAGTGGCCGCTAGT 600
DB 1151 GCGACTGAAATTTGTACAGAGGTTTGGGAAGTAGACCTGGCTAGTGGCCGCTAGT 1210
QY 601 GCGTGAATTTGAGAAAGCCCTCTTGAATGTGTTCCCACTTAAACATCTCATACATA 660
DB 1211 GCGTGAATTTGAGAAAGCCCTCTTGAATGTGTTCCCACTTAAACATCTCATACATA 1270
QY 661 GCGTTCAGATCTGACACAACTGTAAGGCGAGCTTAAGCCACTGTGTGTGAGAGAGCTT 720
DB 1271 GCGTTCAGATCTGACACAACTGTAAGGCGAGCTTAAGCCACTGTGTGTGAGAGAGCTT 1330
QY 721 GAAACCTGAGCACTGAGCACTCAGAAAGTATTCACCAATGATGATGATGGCCATGT 780
DB 1331 GAAACCTGAGCACTGAGCACTCAGAAAGTATTCACCAATGATGATGATGGCCATGT 1390
QY 781 GGTGCTGCTGGGGGGGCTGTGGCCAGCTTTCACAGTTAGCCAGAGGCCAAGCTGCCG 840
DB 1391 GGTGCTGCTGGGGGGGCTGTGGCCAGCTTTCACAGTTAGCCAGAGGCCAAGCTGCCG 1450
QY 841 GGAAGTGAAGACATGTGAAATTTCTGTGTGAAACATGATGAGATGAAATTTACCATTAC 900

DB 1451 GGAAGTGAAGACATGTGAAATTTCTGTGTGAAACATGATGAGATGAAATTTACCATTAC 1510
QY 901 CCAATGTTCTCATCCCCAGAGCAAGTGTGGGTCTGATTAATCTGCAACACAGAGAACGAGA 960
DB 1511 CCAATGTTCTCATCCCCAGAGCAAGTGTGGGTCTGATTAATCTGCAACACAGAGAACGAGA 1570
QY 961 AGAATCTTTTCTCATACAGATCAGAGAGGCTCATACACTGAGGCTGATCATCTCA 1020
DB 1571 AGAATCTTTTCTCATACAGATCAGAGAGGCTCATACACTGAGGCTGATCATCTCA 1630
QY 1021 CCCACACAGACCGGCTTCTCTCCAGTGTGACCTTACACACTGATGCTTACCAAGAT 1080
DB 1631 CCCACACAGACCGGCTTCTCTCCAGTGTGACCTTACACACTGATGCTTACCAAGAT 1690
QY 1081 GATGTTGCCAGAGTCAATGACCAATGTTGCTCCGCCAAGTTCACAGAACTGATTTCTT 1140
DB 1691 GATGTTGCCAGAGTCAATGACCAATGTTGCTCCGCCAAGTTCACAGAACTGATTTCTT 1750
QY 1141 TAAACTAATGACCAATGACCAATGTTGCTCCGCCAAGTTCACAGAACTGATTTCTT 1200
DB 1751 TAAACTAATGACCAATGACCAATGTTGCTCCGCCAAGTTCACAGAACTGATTTCTT 1810
QY 1201 ACACAGCAAGATCCACCTCTGTTCTGTAGCTGACACAGTGAAGTGTGGAACAGAC 1260
DB 1811 ACACAGCAAGATCCACCTCTGTTCTGTAGCTGACACAGTGAAGTGTGGAACAGAC 1870
QY 1261 AGTGACCATCAGACCTTGTGATGAGCGTTTGAATCAACACCTTCAGAAACAAACA 1320
DB 1871 AGTGACCATCAGACCTTGTGATGAGCGTTTGAATCAACACCTTCAGAAACAAACA 1930
QY 1321 CCATTCAGTGTCTGTAGCCCTTAATTTAAGCTTTTCAAGAAAGTGTGGAAGTTTGG 1380
DB 1931 CCATTCAGTGTCTGTAGCCCTTAATTTAAGCTTTTCAAGAAAGTGTGGAAGTTTGG 1990
QY 1381 TAGATAGTGAAGAGGGGGCATCACTGAGAAAGAGCTGATTTGATTCAGGTTGAA 1440
DB 1991 TAGATAGTGAAGAGGGGGCATCACTGAGAAAGAGCTGATTTGATTCAGGTTGAA 2050
QY 1441 AAGAATTAATCTGAATATTTTGTAGGCAAGTCAAGAAAGAACTGTCACCCAAAGC 1500
DB 2051 AAGAATTAATCTGAATATTTTGTAGGCAAGTCAAGAAAGAACTGTCACCCAAAGC 2110
QY 1501 AACTGTAACTCAGAAATTAATTAAGTACAGAAATTAAGTACAGAAATTAAGAAAGAT 1560
DB 2111 AACTGTAACTCAGAAATTAATTAAGTACAGAAATTAAGTACAGAAATTAAGAAAGAT 2170
QY 1561 GGTATATGAACCCCATATACCTTCTGATGATTCACCAATGTTTAACTTTTTC 1620
DB 2171 GGTATATGAACCCCATATACCTTCTGATGATTCACCAATGTTTAACTTTTTC 2230
QY 1621 CTCTCAGCTATCCTCTAATTTCTCTAATTTCAATTTGTTAATTAATTAATTAATTA 1680
DB 2231 CTCTCAGCTATCCTCTAATTTCTCTAATTTCAATTTGTTAATTAATTAATTAATTA 2290
QY 1681 CAATAAGGCAATCTGTGAGAAATTTGAAGCAATTTAGAAATCTTTTGAATTTTCTG 1740
DB 2291 CAATAAGGCAATCTGTGAGAAATTTGAAGCAATTTAGAAATCTTTTGAATTTTCTG 2350
QY 1741 TGGTTTATGGAATTAATGAGGTTTAACTGTGGGGTGAAGGAGAGAGCTTACTCATTTG 1800
DB 2351 TGGTTTATGGAATTAATGAGGTTTAACTGTGGGGTGAAGGAGAGAGCTTACTCATTTG 2410
QY 1801 ACCAGATTTGTTGCTTAACATCCGAGAGATGATTTTGTACAGAAATTAATTAATTA 1860
DB 2411 ACCAGATTTGTTGCTTAACATCCGAGAGATGATTTTGTACAGAAATTAATTAATTA 2470
QY 1861 ATTAATTTTCAAGATATTTTCTCTCAATTAAGTAACTTAATTAATTA 1910
DB 2471 ATTAATTTTCAAGATATTTTCTCTCAATTAAGTAACTTAATTAATTA 2520

RESULT 10

US-09-699-998-10797
Sequence 10797, Application US/09699998
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.
APPLICANT: Geating, David P.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE OF INVENTION: THEREFOR
FILE REFERENCE: 1600,2008-001
CURRENT APPLICATION NUMBER: US/09/699,998
CURRENT FILING DATE: 2000-10-30
PRIOR APPLICATION NUMBER: 60/162,362
PRIOR FILING DATE: 1999-10-29
NUMBER OF SEQ ID NOS: 10905
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10797
LENGTH: 2541
TYPE: DNA
ORGANISM: Homo sapiens
US-09-699-998-10797
Query Match 100.0%; Score 1910; DB 29; Length 2541;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGGGCTGATGCTGACCATGAGATGAGCTCCGCCCGAAGACCGGGTGAGGC 60
DB 611 CTGGGCTGATGCTGACCATGAGATGAGCTCCGCCCGAAGACCGGGTGAGGC 670
QY 61 TCTCTCCAGTGGGTAGTGGGTAGAGTGAAGACATCCACCCCGTGGTACTT 120
DB 671 TCTCTCCAGTGGGTAGTGGGTAGAGTGAAGACATCCACCCCGTGGTACTT 730
QY 121 CCGCTGTGAGTGGATATTCGGAATGGCATTCATTACTGTAGAGAGCAACATTGA 180
DB 731 CCGCTGTGAGTGGATATTCGGAATGGCATTCATTACTGTAGAGAGCAACATTGA 790
QY 181 ACATGCTTACTCTCTATTAACAGATATACGCTCTTATGAGAAATACCAAAACA 240
DB 791 ACATGCTTACTCTCTATTAACAGATATACGCTCTTATGAGAAATACCAAAACA 850
QY 241 TCGAGATTACAAATCTGTCTGCTGCTGCTGAAAGAAAGACACAGTAAAGAAATTAAGA 300
DB 851 TCGAGATTACAAATCTGTCTGCTGCTGCTGAAAGAAAGACACAGTAAAGAAATTAAGA 910
QY 301 GATTGCAATTTCCCAAGCAAGAGCTGAAGGCAAGCTGTAAAGATATCCAAAGA 360
DB 911 GATTGCAATTTCCCAAGCAAGAGCTGAAGGCAAGCTGTAAAGATATCCAAAGA 970
QY 361 ATATACGAATATATGAGAAAGAAAGAAAGAGAGAGATTTGGCCCGAACAATGGC 420
DB 971 ATATACGAATATATGAGAAAGAAAGAAAGAGAGATTTGGCCCGAACAATGGC 1030
QY 421 CATCCAGCAAGAGCTGGAAGAAAGAAAGAGAGAGATTTGGCCCGAACAATGGC 480
DB 1031 CATCCAGCAAGAGCTGGAAGAAAGAAAGAGAGAGATTTGGCCCGAACAATGGC 1090
QY 481 GGAACAGGAAGAGTTCATGCTGCGAGAGATGATCCGGAACAGAGAGCTAGAAAAAGA 540
DB 1091 GGAACAGGAAGAGTTCATGCTGCGAGAGATGATCCGGAACAGAGAGCTAGAAAAAGA 1150
QY 541 GCGACTGAAATTTGACAGAGTTTGGAGAGTAGACCTGAGCTAGTGGCCCGCTAGT 600
DB 1151 GCGACTGAAATTTGACAGAGTTTGGAGAGTAGACCTGAGCTAGTGGCCCGCTAGT 1210
QY 601 GCGTGAATTTGAGAGAGCTCTTATGATGTTTCCCACTTAACATTCATCCATACA 660
DB 1211 GCGTGAATTTGAGAGAGCTCTTATGATGTTTCCCACTTAAACATTCATCCATACA 1270
QY 661 GCGTTCAGACTGTACACAACTGTAAAGGCACTAAGCACTGTGTGAGACAGTCTCT 720
DB 1271 GCGTTCAGACTGTACACAACTGTAAAGGCACTAAGCACTGTGTGAGACAGTCTCT 1330
QY 721 GAAACCTGAGCACTGAGCAACTCAAGAAATTTCCCAATTCATGATGATTTGGCCCATGT 780

DB 1331 GAAACCTGAGCACTGAGCAACTCAAGAAATTTCCCAATTCATGATGATTTGGCCCATGT 1390
QY 781 GGTGTGCTGGGGGGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTSCCAACTGCCCC 840
DB 1391 GGTGTGCTGGGGGGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTSCCAACTGCCCC 1450
QY 841 GGGAGTGGAGCATGTGAAATTTCTGTGAGAAATCTATGAGAAATGAATTTTACATTAC 900
DB 1451 GGGAGTGGAGCATGTGAAATTTCTGTGAGAAATCTATGAGAAATGAATTTTACATTAC 1510
QY 901 CCAATTTCTCATCCCAAGCAAGTGTGGGTGATTAATGCAACAGAGAGAGAGA 960
DB 1511 CCAATTTCTCATCCCAAGCAAGTGTGGGTGATTAATGCAACAGAGAGAGAGA 1570
QY 961 AGAATTTTCTCATTAAGATTCAGAGAGGCTTCATCAACTGAGCTGATTAATCA 1020
DB 1571 AGAATTTTCTCATTAAGATTCAGAGAGGCTTCATCAACTGAGCTGATTAATCA 1630
QY 1021 CCCCAACAGAGCCGCTTTCTCTCAGATGTGACCTACACTAGCTCTTCCAGAT 1080
DB 1631 CCCCAACAGAGCCGCTTTCTCTCAGATGTGACCTACACTAGCTCTTCCAGAT 1690
QY 1081 GATGTTGCAAGTCAATGATGATGTTGCTCCCAAGTTCAGAGAACTGATTTCTT 1140
DB 1691 GATGTTGCAAGTCAATGATGATGTTGCTCCCAAGTTCAGAGAACTGATTTCTT 1750
QY 1141 TAACTAACTGACATGAGCTAGAGAGATTTCTCTGTGTGCCAAGAAAGATTTCTATCC 1200
DB 1751 TAACTAACTGACATGAGCTAGAGAGATTTCTCTGTGTGCCAAGAAAGATTTCTATCC 1810
QY 1201 ACACAGCAAGATCACCTGTTCTGTAGAGTGCAGCACTGAGTGTGAGCAGAGC 1260
DB 1811 ACACAGCAAGATCACCTGTTCTGTAGAGTGCAGCACTGAGTGTGAGCAGAGC 1870
QY 1261 AGTGAACATCAAGACCTTGTGATGAGTGTGAGTCCACACTTCCAGAACACAAA 1320
DB 1871 AGTGAACATCAAGACCTTGTGATGAGTGTGAGTCCACACTTCCAGAACACAAA 1930
QY 1321 CCATATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1380
DB 1931 CCATATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1990
QY 1381 TAGATAGTGAAGAGGGGGGATCACTGAGAGAGAGTGTGATGATTTGATTTGAGT 1440
DB 1991 TAGATAGTGAAGAGGGGGGATCACTGAGAGAGAGTGTGATGATTTGATTTGAGT 2050
QY 1441 AAGAAATTAATGAAATTTTATTTTATGAGCACTGAGAGAGAGAGAGAGAGAGAG 1500
DB 2051 AAGAAATTAATGAAATTTTATTTTATGAGCACTGAGAGAGAGAGAGAGAGAGAG 2110
QY 1501 AACTGTAATCAGAAATTAAGTAACTGAGAAATTAAGTAACTGAGAAATTAAGAAAGAT 1560
DB 2111 AACTGTAATCAGAAATTAAGTAACTGAGAAATTAAGTAACTGAGAAATTAAGAAAGAT 2170
QY 1561 GGTATATGAAACCCCATATACCTTCTTGTGATTAACAAATTTTATTTTTC 1620
DB 2171 GGTATATGAAACCCCATATACCTTCTTGTGATTAACAAATTTTATTTTTC 2230
QY 1621 CTCTAGCAATCTCTCTATTTCTCTATTTCTCTATTTCTCTATTTCTCTATTTCT 1680
DB 2231 CTCTAGCAATCTCTCTATTTCTCTATTTCTCTATTTCTCTATTTCTCTATTTCT 2290
QY 1681 CAATAGAGGCACTGTGAGAGAAATTTGAGAGCAATTTGAGAAATTTTGGATTTCTGT 1740
DB 2291 CAATAGAGGCACTGTGAGAGAAATTTGAGAGCAATTTGAGAAATTTTGGATTTCTGT 2350
QY 1741 TGGTTATGCAATATGAGTGTGATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGT 1800
DB 2351 TGGTTATGCAATATGAGTGTGATGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGT 2410
QY 1801 ACCAGATTTGTTGGCTAACACATCCGGAAGATGATTTTGTGAGAAATTTATTTATTA 1860

```
Db 2411 ACCGATGTTGGCTACACATCCCGAAGATGATTTGTACGAATTAATGTATTTA 2470
QY 1861 ATAAATATTCAGATATTTTCTCTACATTAAGTAACATTAACTTA 1910
Db 2471 ATAAATATTCAGATATTTTCTCTACATTAAGTAACATTAACTTA 2520

RESULT 11
US-09-726-810-2597
; Sequence 2597, Application US/09726810
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2016-001
; CURRENT APPLICATION NUMBER: US/09/726,810
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/168,017
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 3398
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2597
; LENGTH: 2541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-810-2597

Query Match 100.0%; Score 1910; DB 31; Length 2541;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGGTCTCTGATGTCCTGACCATGAGATGTAGAGCTCCCGCCGAAACCGGGTGAAGGC 60
Db 611 CTTGGTCTCTGATGTCCTGACCATGAGATGTAGAGCTCCCGCCGAAACCGGGTGAAGGC 670
QY 61 TCTCTCCACGCTGGGTAGTGGGTAGAGTGAATGAAGACATTCACCCCGTGGTACTT 120
Db 671 TCTCTCCACGCTGGGTAGTGGGTAGAGTGAATGAAGACATTCACCCCGTGGTACTT 730
QY 121 CCGCTCTGAGATGAGATTAATCCGAATGGCATCCATTACTCTGAGAGAGCAACATTGA 180
Db 731 CCGCTCTGAGATGAGATTAATCCGAATGGCATCCATTACTCTGAGAGAGCAACATTGA 790
QY 181 ACATGCTCTCATCTCTATATACAGTATATACAGCTCTTTATGAGAACTAACCAACA 240
Db 791 ACATGCTCTCATCTCTATATACAGTATATACAGCTCTTTATGAGAACTAACCAACA 850
QY 241 TCGAGATTACAAATCTGCTGTCATTCCTGAAAGAAAGACAGTAAGAATTTAAAGA 300
Db 851 TCGAGATTACAAATCTGCTGTCATTCCTGAAAGAAAGACAGTAAGAATTTAAAGA 910
QY 301 GATTGCTATTCCTCCAAAGCAGAAAGAGCTGAAGCGAGCTGTAAACGATATACCAAGA 360
Db 911 GATTGCTATTCCTCCAAAGCAGAAAGAGCTGAAGCGAGCTGTAAACGATATACCAAGA 970
QY 361 ATATACGAATATATATGAGAAAGAGAGAGAGAGATTTGGCCCGGAACATGGC 420
Db 971 ATATACGAATATATATGAGAAAGAGAGAGAGAGATTTGGCCCGGAACATGGC 1030
QY 421 CATTCAGCAAGAGCTGGAAGAAAGAAAGAGAGGATGCAACACAGAGCAGACGCAAT 480
Db 1031 CATTCAGCAAGAGCTGGAAGAAAGAAAGAGAGGATGCAACACAGAGCAGACGCAAT 1090
QY 481 GGAACAGAAACAGTTCATGCTTCGAGAGATGATTCGGAACAGAGAGCTAGAAAAAG 540
Db 1091 GGAACAGAAACAGTTCATGCTTCGAGAGATGATTCGGAACAGAGAGCTAGAAAAAG 1150
QY 541 GCGACTAAAAATTTGACAGAGATTTGGGAAAGTAGACCTGGCTAGAGTGGCCGCTAGT 600
Db 1151 GCGACTAAAAATTTGACAGAGATTTGGGAAAGTAGACCTGGCTAGAGTGGCCGCTAGT 1210
QY 601 GCGTGAAGTGGAGAGCCCTCTAGATGTGTCCCACTTAAGATCTCATCCATACA 660
Db 1211 GCGTGAAGTGGAGAGCCCTCTAGATGTGTCCCACTTAAGATCTCATCCATACA 1270
QY 661 GCGTGAAGTGGAGAGCCCTCTAGATGTGTCCCACTTAAGATCTCATCCATACA 1330
Db 1271 GCGTGAAGTGGAGAGCCCTCTAGATGTGTCCCACTTAAGATCTCATCCATACA 1390
QY 721 GAAACCTGAGCAGCTAGCAACTCAGAAAGTATTCACCAATGATGATTTGGCCATGT 780
Db 1331 GAAACCTGAGCAGCTAGCAACTCAGAAAGTATTCACCAATGATGATTTGGCCATGT 1390
QY 781 GGTGGTCTGGGCGGCTGTGCCACAGTTTCTCAGTTAGCCAGTGGCAACATGCCCCG 840
Db 1391 GGTGGTCTGGGCGGCTGTGCCACAGTTTCTCAGTTAGCCAGTGGCAACATGCCCCG 1450
QY 841 GGGAGTGGAGACATGAGAAATCTCTGTGAGAAACATGAGAGAAATTTACATTAC 900
Db 1451 GGGAGTGGAGACATGAGAAATCTCTGTGAGAAACATGAGAGAAATTTACATTAC 1510
QY 901 CCATGTTCTCATCCCAAGCAAAAGTGGTGTGATTACTGCAACACAGAGACAGAGA 960
Db 1511 CCATGTTCTCATCCCAAGCAAAAGTGGTGTGATTACTGCAACACAGAGACAGAGA 1570
QY 961 AGAAGCTTTTCTCATCAGAGATCAGAGAGGCTCATCATCTGGGCTGATTCATCTCA 1020
Db 1571 AGAAGCTTTTCTCATCAGAGATCAGAGAGGCTCATCATCTGGGCTGATTCATCTCA 1630
QY 1021 CCCACACAGACCGGCTTCTCTCCAGTCCAGCTACACATCACTGCTCTTACAGAT 1080
Db 1631 CCCACACAGACCGGCTTCTCTCCAGTCCAGCTACACATCACTGCTCTTACAGAT 1690
QY 1081 GATGTTCCAGAGTCAATGATGCTGATTTGCTCCCCAAGTTCCAGAAACATGATCTT 1140
Db 1691 GATGTTCCAGAGTCAATGATGCTGATTTGCTCCCCAAGTTCCAGAAACATGATCTT 1750
QY 1141 TAAACTTAATGACATGAGCTAGAGAGATTTCTTCGTGGCCAGAAAGATTTTCATGC 1200
Db 1751 TAAACTTAATGACATGAGCTAGAGAGATTTCTTCGTGGCCAGAAAGATTTTCATGC 1810
QY 1201 ACAAGAGAGATTCACCTCTGTCTGTAGCTGACGCCAGTACTGTTTGAACAGAGC 1260
Db 1811 ACAAGAGAGATTCACCTCTGTCTGTAGCTGACGCCAGTACTGTTTGAACAGAGC 1870
QY 1261 AGTGACCATCAGACGCTCGATGAGCGCTTGAAGTCCAACTTCGAAACCAACAAA 1320
Db 1871 AGTGACCATCAGACGCTCGATGAGCGCTTGAAGTCCAACTTCGAAACCAACAAA 1930
QY 1321 CCATATCAGTGTATGCTAGCCCTTAATTTAGCTTTCTAGAAAGCTTTGAAAGTTTGG 1380
Db 1931 CCATATCAGTGTATGCTAGCCCTTAATTTAGCTTTCTAGAAAGCTTTGAAAGTTTGG 1990
QY 1381 TAGATAGTGAAGGGGGGCACTCACTGAGAAAGCTGATTTTGTATTTCAAGTTTGA 1440
Db 1991 TAGATAGTGAAGGGGGGCACTCACTGAGAAAGCTGATTTTGTATTTCAAGTTTGA 2050
QY 1441 AAGAAATTAATGAATATTTTGTAGCAAGTCAAGAAAGAACTGGTCAACCCAAAAGC 1500
Db 2051 AAGAAATTAATGAATATTTTGTAGCAAGTCAAGAAAGAACTGGTCAACCCAAAAGC 2110
QY 1501 AACTGTAATCAGAAATTAAGTTACTCAGAAATTAAGTGTAGAGTAAAGAAAGAT 1560
Db 2111 AACTGTAATCAGAAATTAAGTTACTCAGAAATTAAGTGTAGAGTAAAGAAAGAT 2170
QY 1561 GGTATATGAAGCCCATATACCTTCCTCTGAGTATACCAATGTATTAACATTTTTC 1620
Db 2171 GGTATATGAAGCCCATATACCTTCCTCTGAGTATACCAATGTATTAACATTTTTC 2230
QY 1621 CTCTGAGTATCTCTTAATTTCTCTTAATTTCAATTTGTTATATTAATCTCTGGCT 1680
Db 2231 CTCTGAGTATCTCTTAATTTCTCTTAATTTCAATTTGTTATATTAATCTCTGGCT 2290
QY 1681 CATTAAGGCACTCTGTGAGAAATTTGAGAGCAATTTTGAATTTTCTCTG 1740
```

Db 2291 CAATAAGGAGCATCTGACAGAAATTGGAAAGCCATTAGAAAACTTTTGGATTTCCTG 2350
Cy 1741 TGGTTATGSCAATATGATGAGCTTATTACTGAGGAGGAGACAGCTTACTCCATTG 1800
Db 2351 TGGTTATGSCAATATGATGAGCTTATTACTGAGGAGGAGACAGCTTACTCCATTG 2410
Cy 1801 ACCAGATTGTTGGCTTACACATCCCGAAGATGATTTTGTGAGAAATTATGTTTAA 1860
Db 2411 ACCAGATTGTTGGCTTACACATCCCGAAGATGATTTTGTGAGAAATTATGTTTAA 2470
Cy 1861 ATAAATATTCAGAGATATTTTCTCTACATATAAGATTAATACTTA 1910
Db 2471 ATAAATATTCAGAGATATTTTCTCTACATATAAGATTAATACTTA 2520

RESULT 12
US-60-213-362-4025
Sequence 4025, Application US/60213362
GENERAL INFORMATION:
APPLICANT: Morris, Macdonald
APPLICANT: Lal, Preeti
TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
FILE REFERENCE: GX-0016 P
CURRENT APPLICATION NUMBER: US/60/213,362
CURRENT FILING DATE: 2000-06-21
NUMBER OF SEQ ID NOS: 8429
SOFTWARE: PERL Program
SEQ ID NO 4025
LENGTH: 2087
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incycle ID No: 235222.9
US-60-213-362-4025

Query Match 99.9%; Score 1909; DB 77; Length 2087;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 CTTGGCTCGATGTCGTGACCAATGAGATGTGAGCTCCGCCCGAAGACCGGGTGAAGGC 60
Db 179 CTTGGCTCGATGTCGTGACCAATGAGATGTGAGCTCCGCCCGAAGACCGGGTGAAGGC 218
Cy 61 TCTCTCCAGCTGGGTAGTGGGTAGAGTGAATGAGATTCACCCCGTGGTACTT 120
Db 239 TCTCTCCAGCTGGGTAGTGGGTAGAGTGAATGAGATTCACCCCGTGGTACTT 298
Cy 121 CCGCTCGAGATTAGATTATCCGAATGCGATCCATTACTCTGAGGAGGCAACATTGA 180
Db 299 CCGCTCGAGATTAGATTATCCGAATGCGATCCATTACTCTGAGGAGGCAACATTGA 358
Cy 181 ACATGCTTCATCTCTATTAACAAGTATATCAGCTCTTTATGAGAACTTACCAAAACA 240
Db 359 ACATGCTTCATCTCTATTAACAAGTATATCAGCTCTTTATGAGAACTTACCAAAACA 418
Cy 241 TCGAGATTAGCAATCTGCTGCACTTCTGAAAAAGAAAGACAGATTAAGAAATTTAAAGA 300
Db 419 TCGAGATTAGCAATCTGCTGCACTTCTGAAAAAGAAAGACAGATTAAGAAATTTAAAGA 478
Cy 301 GATTGCAATTTCCAAAGAGAGAGCTGAAGGCGAGAGCTGTAAAAAGATATACCAAGA 360
Db 479 GATTGCAATTTCCAAAGAGAGAGCTGAAGGCGAGAGCTGTAAAAAGATATACCAAGA 538
Cy 361 ATATACAGATATATATGAGAAAGAGAGAGAGAGAGAGAGATTTGGCCCGGAAACATGCG 420
Db 539 ATATACAGATATATATGAGAAAGAGAGAGAGAGAGAGAGATTTGGCCCGGAAACATGCG 598
Cy 421 CATCAGCAGAGCTGAGAAAGAGAAACAGAGGGTACCAACAGAGAGAGAGCAATT 480

Db 599 CATCAGCAGAGCTGAGAAAGAGAAACAGAGGGTAGCAACAGAGAGAGAGCAATT 658
Cy 481 GAAACAGAAACAGATTCCATGCTTCGAGAGATATATCCGAGACAGAGAGCTAGAAAGA 540
Db 659 GAAACAGAAACAGATTCCATGCTTCGAGAGATATATCCGAGAGAGAGCTAGAAAGA 718
Cy 541 GCGACTGAAATTTATACAGAGATTGGGAAAGTAGACCTGAGCTAGAGTGGCCGCTAGT 600
Db 719 GCGACTGAAATTTATACAGAGATTGGGAAAGTAGACCTGAGCTAGAGTGGCCGCTAGT 778
Cy 601 GCTGCACTTGAGAAAGCCCTCTTATGATGTGTTCCCACTTAAACATCTCATTCATACA 660
Db 779 GCTGCACTTGAGAAAGCCCTCTTATGATGTGTTCCCACTTAAACATCTCATTCATACA 838
Cy 661 GCTTCAGACTGTACACAACTGTAAGGCCAGTAAAGCACTGAGTGAAGAGAGCTT 720
Db 839 GCTTCAGACTGTACACAACTGTAAGGCCAGTAAAGCACTGAGTGAAGAGAGCTT 898
Cy 721 GAAACCTGAGCACTGAGCAACTCAGAAAGTATTTCCCAATCATGATGATTTGCGCCATGT 780
Db 899 GAAACCTGAGCACTGAGCAACTCAGAAAGTATTTCCCAATCATGATGATTTGCGCCATGT 958
Cy 781 GGTGCTGCTGGGGGGCTGTGSCCAAGTTCTCCAGTTAGCCAGTSCAACAAGTGGCCG 840
Db 959 GGTGCTGCTGGGGGGCTGTGSCCAAGTTCTCCAGTTAGCCAGTSCAACAAGTGGCCG 1018
Cy 841 GCGAGTGAAGACATGTGAATTTCTGTGAAAACTGATGAGAAATTTACATTAC 900
Db 1019 GCGAGTGAAGACATGTGAATTTCTGTGAAAACTGATGAGAAATTTACATTAC 1078
Cy 901 CCATGTTCTCATCCCGCAAGCAAGTGTGGGTCTGATTAATGCAACACAGAGAGAGAGA 960
Db 1079 CCATGTTCTCATCCCGCAAGCAAGTGTGGGTCTGATTAATGCAACACAGAGAGAGAGA 1138
Cy 961 AGAATTTTCTCATATACAGATCAGATCAGAGGCTCATCACTGAGCTGATTCATCTCA 1020
Db 1139 AGAATTTTCTCATATACAGATCAGATCAGAGGCTCATCACTGAGCTGATTCATCTCA 1198
Cy 1021 CCCACACAGACCGCGTTTCTCTCAAGTGTGACCTTACACTCACTGCTTACCAAGAT 1080
Db 1199 CCCACACAGACCGCGTTTCTCTCAAGTGTGACCTTACACTCACTGCTTACCAAGAT 1258
Cy 1081 GATGTTGACAGATGATGAGCATGTTGCTCCCGCAAGTTCAGAGAACTGATTCCTT 1140
Db 1259 GATGTTGACAGATGATGAGCATGTTGCTCCCGCAAGTTCAGAGAACTGATTCCTT 1318
Cy 1141 TAACTTAATCTGACATGAGATAGAGAGATTTCTCTGTGCGCAAGAAAGATTTCAATCC 1200
Db 1319 TAACTTAATCTGACATGAGATAGAGAGATTTCTCTGTGCGCAAGAAAGATTTCAATCC 1378
Cy 1201 ACACAGCAAGATCCACTCTGTTCTGTAGCTGAGCCAGCTGTTGAGCAAGAGC 1260
Db 1379 ACACAGCAAGATCCACTCTGTTCTGTAGCTGAGCCAGCTGTTGAGCAAGAGC 1438
Cy 1261 AGTGACCATCAAGACCTTGGATGAGCGTTTGAGTCAACACCTTCCAAAGAACAAAA 1320
Db 1439 AGTGACCATCAAGACCTTGGATGAGCGTTTGAGTCAACACCTTCCAAAGAACAAAA 1498
Cy 1321 CCATATCAGTGTACTGTGACCCCTTAATTTAAGCTTTTAAAGCTTTGGAAGTTTTG 1380
Db 1499 CCATATCAGTGTACTGTGACCCCTTAATTTAAGCTTTTAAAGCTTTGGAAGTTTTG 1558
Cy 1381 TAGATAGTGAAGAGGGGGCATTCACTGAGAAAGAGCTGATTTGTATTTAGGTTTGA 1440
Db 1559 TAGATAGTGAAGAGGGGGCATTCACTGAGAAAGAGCTGATTTGTATTTAGGTTTGA 1618
Cy 1441 AAGAAATATCTGAACATATTTTATGAGCAAGTCAAGAAAGAGACATGCTCCAAAAAGC 1500
Db 1619 AAGAAATATCTGAACATATTTTATGAGCAAGTCAAGAAAGAGACATGCTCCAAAAAGC 1678
Cy 1501 AACTGTAATCTGAAATTAAGTTACTCAGAAATTAAGAGCTCAGAAATTTAAGAAAGAT 1560
Db 1679 AACTGTAATCTGAAATTAAGTTACTCAGAAATTTAAGAGCTCAGAAATTTAAGAAAGAT 1738

QY 1561 GGTATATGAACCCCATATACCTTCCTTGATTCACCAATGTATTAACATTTTTC 1620
 Db 1739 GGTATATGAACCCCATATACCTTCCTTGATTCACCAATGTATTAACATTTTTC 1798
 QY 1621 CTCACGATCTCTCTCTATTTCTCTCTATTTTCAATTTGTTTATTTTACCTCTGGGCT 1680
 Db 1799 CTCACGATCTCTCTCTATTTCTCTCTATTTTCAATTTGTTTATTTTACCTCTGGGCT 1858
 QY 1681 CAATAGGGCATCTGTGAGAAATTTGAGAGCATTTGAAATCTTTTGATTTTCTG 1740
 Db 1859 CAATAGGGCATCTGTGAGAAATTTGAGAGCATTTGAAATCTTTTGATTTTCTG 1918
 QY 1741 TGGTTTATGGCAATATGATGAGAGCTTATTACTGGGGTGAGGAGACGCTTACTCATTTG 1800
 Db 1919 TGGTTTATGGCAATATGATGAGAGCTTATTACTGGGGTGAGGAGACGCTTACTCATTTG 1978
 QY 1801 ACCAGATGTTTGGTACACATCCCGAAGATATTTTGCAGGAATTTATTTATTTA 1860
 Db 1979 ACCAGATGTTTGGTACACATCCCGAAGATATTTTGCAGGAATTTATTTATTTA 2038
 QY 1861 ATAAATATTTAGATATTTTCTCTACAAATTAAGTAAACATTTACTT 1909
 Db 2039 ATAAATATTTAGATATTTTCTCTACAAATTAAGTAAACATTTACTT 2087

RESULT 13

US-60-278-561-6951
 / Sequence 6951, Application US/60278561
 / GENERAL INFORMATION:
 / APPLICANT: Morris, MacDonald
 / APPLICANT: Lal, Preeti
 / APPLICANT: Deep, Dindi
 / TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
 / TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
 / FILE REFERENCE: GX-0012-1 P
 / CURRENT APPLICATION NUMBER: US/60/278, 561
 / CURRENT FILING DATE: 2001-03-23
 / NUMBER OF SEQ ID NOS: 15598
 / SOFTWARE: PERL Program
 / SEQ ID NO 6951
 / LENGTH: 2087
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURES:
 / NAME/KEY: misc feature
 / OTHER INFORMATION: Incyte ID No: 235222.9
 US-60-278-561-6951

Query Match 99.9%; Score 1909; DB 84; Length 2087;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGTCCTGATGCTTGACCATGAGATGTGAGCTTCCGCCGAGACCGGGTGAAGGC 60
 Db 179 CTGTCCTGATGCTTGACCATGAGATGTGAGCTTCCGCCGAGACCGGGTGAAGGC 238
 QY 61 TCTCTCCAGCTGGGTGAGCGGTGAGAGGTGAATGAAGATTCACCCCGTGGTACTT 120
 Db 239 TCTCTCCAGCTGGGTGAGCGGTGAGAGGTGAATGAAGATTCACCCCGTGGTACTT 298
 QY 121 CCGCTCGAGTGAATTAATTCGAATGCGATTCATTACTCTGAGGAAGCAACATTGA 180
 Db 299 CCGCTCGAGTGAATTAATTCGAATGCGATTCATTACTCTGAGGAAGCAACATTGA 358
 QY 181 ACATGCTTCATCTCTATTAACAAGTATATCAAGCTCTTTATGGAAGTACCAAAACA 240
 Db 359 ACATGCTTCATCTCTATTAACAAGTATATCAAGCTCTTTATGGAAGTACCAAAACA 418
 QY 241 TCGAGATTAACAATCTGCTGCAATCTGAAAAGAAAGACACAGTAAAGAAATTAAGA 300
 Db 419 TCGAGATTAACAATCTGCTGCAATCTGAAAAGAAAGACACAGTAAAGAAATTAAGA 478

QY 301 GATTGATTTCCGAAGACGAAGAGCTGAAGGACGCTGTTAAACGATATACCAAGA 360
 Db 479 GATTGATTTCCGAAGACGAAGAGCTGAAGGACGCTGTTAAACGATATACCAAGA 538
 QY 361 ATATACGAATATATGAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 420
 Db 539 ATATACGAATATATGAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 598
 QY 421 CATCCAGAAAGAGCTGAAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 480
 Db 599 CATCCAGAAAGAGCTGAAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAT 658
 QY 481 GGAACAGAAAGATTCATGCTTCGAGAGATGATCCGGAACAGAGCTAGAAAAGA 540
 Db 659 GGAACAGAAAGATTCATGCTTCGAGAGATGATCCGGAACAGAGCTAGAAAAGA 718
 QY 541 GCGACTGAAATTTGACAGAGTTTGGAAAGTACCCCTGGCTAGTGGCCCGTAGT 600
 Db 719 GCGACTGAAATTTGACAGAGTTTGGAAAGTACCCCTGGCTAGTGGCCCGTAGT 778
 QY 601 GCGTGACTTGGAGAAAGCCCTCCTTATGATGTTTCCCACTTAACAGTCTCATATCA 660
 Db 779 GCGTGACTTGGAGAAAGCCCTCCTTATGATGTTTCCCACTTAACAGTCTCATATCA 838
 QY 661 GCGTTCAGACTGTACACAACTGTAAAGCCAGCTAAGCCACTGTGTGACAGGTCTT 720
 Db 839 GCGTTCAGACTGTACACAACTGTAAAGCCAGCTAAGCCACTGTGTGACAGGTCTT 898
 QY 721 GAAACCTGAGACCTGAGAGAACTCAGAAAGTATTCACAAATCATGATGATTCGCCATGT 780
 Db 899 GAAACCTGAGACCTGAGAGAACTCAGAAAGTATTCACAAATCATGATGATTCGCCATGT 958
 QY 781 GGTGTGCTTGGGCGGCTGTGCCCAGATTTCTCCAGTTAGCAGTGAAGCCACATGCCCCG 840
 Db 959 GGTGTGCTTGGGCGGCTGTGCCCAGATTTCTCCAGTTAGCAGTGAAGCCACATGCCCCG 1018
 QY 841 GGGAGTGGAGACATGTGGAATTTCTGTGGAAGAACTGATGAGAAATTAACCATTAAC 900
 Db 1019 GGGAGTGGAGACATGTGGAATTTCTGTGGAAGAACTGATGAGAAATTAACCATTAAC 1078
 QY 901 CCATGTTTCATCCCAAGCAAAAGTGTGGGCTGTGATTAACGCAACACAGAGAGAGA 960
 Db 1079 CCATGTTTCATCCCAAGCAAAAGTGTGGGCTGTGATTAACGCAACACAGAGAGAGA 1138
 QY 961 AGAATTTTCTCATACAGATACAGAGAGGCTCATACATGAGGCTGATTCATATCA 1020
 Db 1139 AGAATTTTCTCATACAGATACAGAGAGGCTCATACATGAGGCTGATTCATATCA 1198
 QY 1021 CCCACAGACAGCGGCTTCTCCAGTGTGACCTTACACTCACTGCTTTACCAAGT 1080
 Db 1199 CCCACAGACAGCGGCTTCTCCAGTGTGACCTTACACTCACTGCTTTACCAAGT 1258
 QY 1081 GATGTGCGCAGATCAGTATGATTTGTCTCCCAAGTTCAGAAATCAGATTCCTT 1140
 Db 1259 GATGTGCGCAGATCAGTATGATTTGTCTCCCAAGTTCAGAAATCAGATTCCTT 1318
 QY 1141 TAACTTAATGACCATGACATTAAGAGAAATTTCTTCTGTGCGCAAGAAAGATTTCAATCC 1200
 Db 1319 TAACTTAATGACCATGACATTAAGAGAAATTTCTTCTGTGCGCAAGAAAGATTTCAATCC 1378
 QY 1201 ACACAGCAAGATCCACCTCTGTTCTGTAGCTGACAGCAAGTGAATGTTGAGACAGAGC 1260
 Db 1379 ACACAGCAAGATCCACCTCTGTTCTGTAGCTGACAGCAAGTGAATGTTGAGACAGAGC 1438
 QY 1261 AGTACCATCACAGACTTTCATGATGAGCGTTTGAATCCAAACCTTCCAAAGAACAAAAA 1320
 Db 1439 AGTACCATCACAGACTTTCATGATGAGCGTTTGAATCCAAACCTTCCAAAGAACAAAAA 1498
 QY 1321 CCATATCAGTATCAGTGAAGCCCTTAATTTAAGCTTCTAAGAGCTTGAAGATTTTG 1380
 Db 1499 CCATATCAGTATCAGTGAAGCCCTTAATTTAAGCTTCTAAGAGCTTGAAGATTTTG 1558

QY 1381 TAGATAGTAAAGGGGGGATCAGTGAAGAGCTGATTTTGTATTTGAGTTTGA 1440
DB 1559 TAGATAGTAAAGGGGGGATCAGTGAAGAGCTGATTTTGTATTTGAGTTTGA 1618
QY 1441 AAGAAATTAAGTAAATTTTGTAGGCAAGTGAAGAGCAATGTCACCCAAAGC 1500
DB 1619 AAGAAATTAAGTAAATTTTGTAGGCAAGTGAAGAGCAATGTCACCCAAAGC 1678
QY 1501 AACTGTAACTCAGAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 1560
DB 1679 AACTGTAACTCAGAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 1738
QY 1561 GGTATTAATGAAGGGGATATACCTTCTGAGTAACTGAGTAACTGAGTAACT 1620
DB 1739 GGTATTAATGAAGGGGATATACCTTCTGAGTAACTGAGTAACTGAGTAACT 1798
QY 1621 CTCTCAGCTATCTCTTAATTTCTCTTAATTTCTTAATTTCTTAATTTCTTA 1680
DB 1799 CTCTCAGCTATCTCTTAATTTCTCTTAATTTCTTAATTTCTTAATTTCTTA 1858
QY 1681 CAATTAAGGCAATCTGTCAGAAATTTGGAAGCAATTTAGAAATCTTTGAGT 1740
DB 1859 CAATTAAGGCAATCTGTCAGAAATTTGGAAGCAATTTAGAAATCTTTGAGT 1918
QY 1741 TGGTTATGGCAATATGATGAGCTTATTAAGTAAATTAAGTAAATTAAGTAA 1800
DB 1919 TGGTTATGGCAATATGATGAGCTTATTAAGTAAATTAAGTAAATTAAGTAA 1978
QY 1801 ACCAGATTTGTTGCTTAACATCTCCGAGAAATGATTTTGTGAGTAAATTT 1860
DB 1979 ACCAGATTTGTTGCTTAACATCTCCGAGAAATGATTTTGTGAGTAAATTT 2038
QY 1861 AATAATTTTCAGATATTTTCTCTACATTAAGTAAATTAAGTAAATTAAGTAA 1909
DB 2039 AATAATTTTCAGATATTTTCTCTACATTAAGTAAATTAAGTAAATTAAGTAA 2087

RESULT 14

US-60-324-185-16983
Sequence 16983, Application US/60324185
GENERAL INFORMATION:
APPLICANT: Morris, MacDonald
APPLICANT: Lal, Preeti
TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY
FILE REFERENCE: GX-0019-1 P
CURRENT FILING DATE: 2001-09-21
NUMBER OF SEQ ID NOS: 35862
SOFTWARE: PERL Program
SEQ ID NO 16983
LENGTH: 2087
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incycle ID No: 235222.9
US-60-324-185-16983

Query Match 99.9%; Score 1909; DB 89; Length 2087;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGTGCTGATGTCGACATGAGATGTGAGCTCCGCGCGAGAGCCGGGTGAGGC 60
DB 179 CTGTGCTGATGTCGACATGAGATGTGAGCTCCGCGCGAGAGCCGGGTGAGGC 238
QY 61 TCTCTCCAGCTGGTATGCGGTGAGAGTGAATGAAGCACTTCCACCCGTCGTACT 120
DB 239 TCTCTCCAGCTGGTATGCGGTGAGAGTGAATGAAGCACTTCCACCCGTCGTACT 298

QY 121 CCGCTGAGAGTGAATTAATGAGTGCATCCATTTACTCTGAGGAAGCAATTA 180
DB 299 CCGCTGAGAGTGAATTAATGAGTGCATCCATTTACTCTGAGGAAGCAATTA 358
QY 181 ACATGCTTCACTCTTAATTAATGAATATGAGTCTTTATTGAGAAATTAACCA 240
DB 359 ACATGCTTCACTCTTAATTAATGAATATGAGTCTTTATTGAGAAATTAACCA 418
QY 241 TCGAGATTAACAATCTGCTGATTTCTGAGAAAGAGCAAGTAAAGTAAAGTAA 300
DB 419 TCGAGATTAACAATCTGCTGATTTCTGAGAAAGAGCAAGTAAAGTAAAGTAA 478
QY 301 GATTGCAATTTCCCAAGCAGAGAGTGAAGGCAAGCTGTTAAACGATATACCAAGA 360
DB 479 GATTGCAATTTCCCAAGCAGAGAGTGAAGGCAAGCTGTTAAACGATATACCAAGA 538
QY 361 AATATCAGAAATTAATGAAGAAAGAAAGAGAGAGAGTGGCCGGAACATGAGC 420
DB 539 AATATCAGAAATTAATGAAGAAAGAAAGAGAGAGTGGCCGGAACATGAGC 598
QY 421 CATCAGCAAGAGCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 599 CATCAGCAAGAGCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 658
QY 481 GGAACAGAGACAGTTCCATGCTTTCGAGAGATGATCCGGAACAGAGAGCTAGAAAGA 540
DB 659 GGAACAGAGACAGTTCCATGCTTTCGAGAGATGATCCGGAACAGAGAGCTAGAAAGA 718
QY 541 GCGACTGAAATTTGACAGAGTTGGAGAGTGAAGCTTGGCTGAGTGGCCGCTAGT 600
DB 719 GCGACTGAAATTTGACAGAGTTGGAGAGTGAAGCTTGGCTGAGTGGCCGCTAGT 778
QY 601 GCTGACTTGAAGAGCCCTCTTATGATGTTTCCCACTTAACAGTCTCATCATACA 660
DB 779 GCTGACTTGAAGAGCCCTCTTATGATGTTTCCCACTTAACAGTCTCATCATACA 838
QY 661 GCTTCAACCTGTACACAACTGTAAGGCGAGTACAGCTGTTGAGAGAGTCTT 720
DB 839 GCTTCAACCTGTACACAACTGTAAGGCGAGTACAGCTGTTGAGAGAGTCTT 898
QY 721 GAAACCTGAGAGCTGAGCAACTGAGAAATTTCCCAATGATGATGATGCGCATGT 780
DB 899 GAAACCTGAGAGCTGAGCAACTGAGAAATTTCCCAATGATGATGATGCGCATGT 958
QY 781 GGTGCTGCTGGGCGGCTGTGCTCCACAGTTTCTCAGTATGAGTGCACACAGTCCG 840
DB 959 GGTGCTGCTGGGCGGCTGTGCTCCACAGTTTCTCAGTATGAGTGCACACAGTCCG 1018
QY 841 GGAAGTGAAGCATGTGAATTTCTGTGAGAAACTGATGAGATGAATTTACCATTA 900
DB 1019 GGAAGTGAAGCATGTGAATTTCTGTGAGAAACTGATGAGATGAATTTACCATTA 1078
QY 901 CCATGTTTCATCCCAAGCAAGAGTGTGATTTAGTGAACACAGAGAAAGAGAG 960
DB 1079 CCATGTTTCATCCCAAGCAAGAGTGTGATTTAGTGAACACAGAGAAAGAGAG 1138
QY 961 AGAATTTTCCCATACAGATCAGCAGAGGCTCATCACTGAGGCTGATTTCACTCA 1020
DB 1139 AGAATTTTCCCATACAGATCAGCAGAGGCTCATCACTGAGGCTGATTTCACTCA 1198
QY 1021 CCCACACAGAGCGGCTTTCTGTCAGTGTGACCTACACACTGCTCTTTACAGAT 1080
DB 1199 CCCACACAGAGCGGCTTTCTGTCAGTGTGACCTACACACTGCTCTTTACAGAT 1258
QY 1081 GATGTTGCGAGAGTGAAGTGAATTTGCTGCTCCCAAGTTCAGAGAACTGATTTCT 1140
DB 1259 GATGTTGCGAGAGTGAAGTGAATTTGCTGCTCCCAAGTTCAGAGAACTGATTTCT 1318
QY 1141 TAAACTAATGACATGACATGAGAGAGATTTCTGCTGCTGCGCAAGAGATTTCACTC 1200
DB 1319 TAAACTAATGACATGACATGAGAGAGATTTCTGCTGCTGCGCAAGAGATTTCACTC 1378
QY 1201 ACAACAGAAAGATCCACTCTGTTCTGATGCTGACAGCAAGTGTGTTGAGACAGAGC 1260

QY	1141	TAACTACTAGTACCATGAGCTAGAGAGAGATTTCTTCCTGTCGCGACGAAAGATTTGATCC	1250
Db	1319	TAAACTAATCTGACATGAGCTAGAGAGAGATTTCTTCCTGTCGCGAAAGATTTGATCC	1378
QY	1201	ACACAGCAGAAGATCCACTCTGTTCCTGAGCTGACGACACGTGACTGTGTGACAGAGC	1260
Db	1319	ACACAGCAGAAGATCCACTCTGTTCCTGAGCTGACGACACGTGACTGTGTGACAGAGC	1438
QY	1261	AGTGCATCATCAGACCTTCGATGAGCGTTTGAATCCAAACCTTCCAGAACACAAA	1320
Db	1439	AGTGACCATCATCAGACCTTCGATGAGCGTTTGAATCCAAACCTTCCAGAACACAAA	1498
QY	1331	CCATATCATGRTACTCTGTAGCCCTTAAATTTAAGCTTCTAAGAAAGCTTTGAGATTTTG	1380
Db	1499	CCATATCATGRTACTCTGTAGCCCTTAAATTTAAGCTTCTAAGAAAGCTTTGAGATTTTG	1558
QY	1381	TAGATATGAGAGAAAGGGGGGCATCACTGAGAAAGAGTGATTTTGTATTTCCAGTTTGA	1440
Db	1559	TAGATATGAGAGAAAGGGGGGCATCACTGAGAAAGAGTGATTTTGTATTTCCAGTTTGA	1618
QY	1441	AAGAAATATCTGAAACATATTTTTTTGTAGSCAAGTCAGAAAGAGACATGTCACCCAAAAGC	1500
Db	1619	AAGAAATATCTGAAACATATTTTTTTGTAGSCAAGTCAGAAAGAGACATGTCACCCAAAAGC	1678
QY	1501	AACCTTAATCTAGAGAAATTAAGTACTCTGAGAAATTAAGAGCTCAAGAAATTAAGAAAGAT	1560
Db	1679	AACCTTAATCTAGAGAAATTAAGTACTCTGAGAAATTAAGAGCTCAAGAAATTAAGAAAGAT	1738
QY	1561	GGTATTAATGAACCCCATATACCCCTCTCTGTGATTCACCAATTTGTTAAACATTTTTTC	1620
Db	1739	GGTATTAATGAACCCCATATACCCCTCTCTGTGATTCACCAATTTGTTAAACATTTTTTC	1798
QY	1621	CTCTCAGCTATTCCTTAATTTCTCTTAATTTCAATTTGTTTATTTTAACTCTGGGCT	1680
Db	1799	CTCTCAGCTATTCCTTAATTTCTCTTAATTTCAATTTGTTTATTTTAACTCTGGGCT	1858
QY	1681	CAATTAAGGGCATTCTGTGAGAAATTTGGAAGCATTTGAAGAAATCTTTGATTTTCTG	1740
Db	1859	CAATTAAGGGCATTCTGTGAGAAATTTGGAAGCATTTGAAGAAATCTTTGATTTTCTG	1918
QY	1741	TGGTTTATGCGCAATATGATATGATGAGCTTTTACTGGGGTGAGGAGCAGCTTACTCCATTTG	1800
Db	1919	TGGTTTATGCGCAATATGATATGATGAGCTTTTACTGGGGTGAGGAGCAGCTTAYTCCATTTG	1978
QY	1801	ACCAGATTTGTTGCTAAACAATCCCGAAGAAATGATTTTGTGAGAAATTTGTTATTTTA	1860
Db	1979	ACCAGATTTGTTGCTAAACAATCCCGAAGAAATGATTTTGTGAGAAATTTGTTATTTTA	2038
QY	1861	ATTAATATTTTCAGATATTTTTTCCCTCAATTAAGTAAGTAACAATTAACCTTA	1910
Db	2039	ATTAATATTTTCAGATATTTTTTCCCTCAATTAAGTAAGTAACAATTAACCTTA	2088

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2004, 09:58:38 ; Search time 759 Seconds

(without alignments)
10690.459 Million cell updates/sec

Title: US-09-831-452-2

Perfect score: 1910
Sequence: 1 ctgtgctctgctgctgacc.....ataagtaacatuaactta 1910

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002s:*

7: Geneseq2003as:*

8: Geneseq2003bs:*

9: Geneseq2003cs:*

10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1910	100.0	1910	3	AAAS9852 Human AMS
2	1910	100.0	2003	5	AAAF57478 Human SAP
3	1043	54.6	1043	3	AAAG9709 Human ova
4	1043	54.6	1043	6	ABN72603 Ovarian c
5	1043	54.6	1043	8	ADA08768 Human ova
6	1040	54.5	1041	3	AAAG9708 Human ova
7	1040	54.5	1041	6	ABN72602 Ovarian c
8	1040	54.5	1041	8	ADA08767 Human ova
9	879.8	46.1	1384	3	AAAS9853 Murine AM
10	682	35.7	1236	5	AAAS9852 Human encod
11	634.4	33.2	1236	5	AAAS9852 Human encod
12	682	22.2	1621	7	ABZ56929 Human AMS
13	418.2	21.9	1392	7	ABZ56931 Mouse AMS
14	418.2	21.9	1428	7	ABZ56930 Mouse AMS
15	413	21.6	4052	4	AAIS9567 Human pol
16	413	21.6	4714	6	ABK92424 Human pol
17	407	20.9	1973	6	ABK13102 Human sec
18	398.6	20.9	407	8	ACH47766 Human inf
19	349.8	18.3	365	5	ABAI6038 Human ner
20	301	15.8	1318	4	AAI61352 Human inf
21	295.4	15.5	406	8	ACH47349 Human inf
22	286	15.0	348	8	ACH31362 Human bon
23	281	14.7	431	8	ACH47326 Human inf

24	262.4	13.7	1898	9	ADB61940	Abd61940 Human cDN
25	245.2	12.8	277	3	AAC02840	Aac02840 Human sec
26	233.8	12.2	301	4	AAFP92297	AAFP92297 Bovine ma
27	183.8	9.6	1611	4	ABL03899	AbI03899 Drosophi
28	125.4	6.6	730	2	AAZ15641	Aaz15641 Human gen
29	118.6	6.2	775	4	AAH33413	Aah33413 Human col
30	115.4	6.0	1184	4	AAH06901	Aah06901 Human rep
31	115.4	6.0	1184	4	ABA08120	Abao8120 Human ova
32	114.4	6.0	1877	3	AAAC39460	Aac39460 Arabidops
33	113.6	6.0	386	4	AAAS98982	Aas98982 Novel hum
34	100.8	5.3	387	5	AAAF67287	AAf67287 Novel hum
35	93	4.9	1010	3	AAAC40578	Aac40578 Arabidops
36	88.4	4.6	1025	3	AAAC46416	Aac46416 Arabidops
37	68.4	3.6	4331	4	ABU03898	AbU03898 Drosophi
38	60	3.1	60	6	ABN39714	Abn39714 Human spl
39	56.8	3.0	591	2	AAAX20389	Aax20389 Borrelia
40	55.2	2.9	575	4	AAI23315	AAI23315 Probe #13
41	55.2	2.9	575	4	ABA68422	Abae68422 Human foe
42	55.2	2.9	575	4	AAI48636	AAI48636 Probe #17
43	55.2	2.9	575	4	ABAS0472	Abas0472 Human bre
44	55.2	2.9	575	4	ABAS5413	Abas5413 Probe #13
45	55.2	2.9	575	4	AAK42567	Aak42567 Human bon

ALIGNMENTS

RESULT 1
ID AAAS9852 standard; cDNA; 1910 BP.
AC AAAS9852;
DT 13-OCT-2000 (first entry)
DE Human AMSH nucleotide sequence.
KM AMSH; signal transduction; SH3 domain; cytokine; STAM; severe infection;
KW cancer; autoimmune disorder; human; ss.
XX
OS Homo sapiens.
XX
PN WO200029436-A1.
XX
PD 25-MAY-2000.
XX
PF 12-NOV-1999; 99WO-JP006309.
XX
PR 12-NOV-1998; 98JP-00322674.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Sugamura K, Tanaka N;
XX
DR WPI; 2000-387743/33.
XX
PT P-FSDB; AAB03791.
XX
PT Protein AMSH being signal transducer interacting with SH3 domain of
PT cytokine signal transducer STAM, for diagnosis and development of drugs
PT of treating diseases e.g. severe infection, cancers and autoimmune
PT disorders.
XX
PS Claim 4; Page 14-15; 22pp; Japanese.
XX
PS This sequence represents the human AMSH gene. The AMSH protein is
XX involved in signal transduction, and interacts with the SH3 domain of
XX cytokine signal transducer STAM. The invention includes human and murine
XX AMSH protein and gene sequences (hAMSH and mAMSH), and also includes
XX antibodies directed against hAMSH and mAMSH. The AMSH protein can be used
XX in the diagnosis and development of drugs for treating diseases due to
XX the dysfunction of cytokine signal transduction e.g. severe infection,
XX cancers and autoimmune disorders

SQ Sequence 1910 BP; 578 A; 424 C; 421 G; 487 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1910; DB 3; Length 1910;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGTCCTGATGCTGACCATGAGATGAGCTTCCCGCCGGAAGACCGGGGTGAGGGC 60
 DB 1 CTTGTCCTGATGCTGACCATGAGATGAGCTTCCCGCCGGAAGACCGGGGTGAGGGC 60
 QY 61 TCTCTCCAGCTGGGTAGTGGGTAGAGGTGAATGAAGATTCACCCCGTGGTACTT 120
 DB 61 TCTCTCCAGCTGGGTAGTGGGTAGAGGTGAATGAAGATTCACCCCGTGGTACTT 120
 QY 121 CCGCTCTGAGATTGATATCCGAATGGCATTCATTACTCTGAGGAGGCAACATTGA 180
 DB 121 CCGCTCTGAGATTGATATCCGAATGGCATTCATTACTCTGAGGAGGCAACATTGA 180
 QY 181 ACATGCTTCTATCTCTATTAACAAGTATATCAGCTCTTTATTAAGAACTACCAAAAG 240
 DB 181 ACATGCTTCTATCTCTATTAACAAGTATATCAGCTCTTTATTAAGAACTACCAAAAG 240
 QY 241 TCGAGATTACAAATCTCTGTCATTCCTGAAAAGAAAGACAGTAAGAAATTTAAAG 300
 DB 241 TCGAGATTACAAATCTCTGTCATTCCTGAAAAGAAAGACAGTAAGAAATTTAAAG 300
 QY 301 GATTGCAATTTCCCAAGACAGAGCTGAAGCGAGAGCTGTAAAGATATACCAAGA 360
 DB 301 GATTGCAATTTCCCAAGACAGAGCTGAAGCGAGAGCTGTAAAGATATACCAAGA 360
 QY 361 ATATACAGATATATATGAAGAAAGAAAGAAAGAGAGCAAGAAATTTGGCCGGAACAT 420
 DB 361 ATATACAGATATATATGAAGAAAGAAAGAAAGAGAGCAAGAAATTTGGCCGGAACAT 420
 QY 421 CATTCAGACAGAGCTGGAAGAAAGAAAGAGAGCTGACACAGAGAGCAAGCAATT 480
 DB 421 CATTCAGACAGAGCTGGAAGAAAGAAAGAGAGCTGACACAGAGAGCAAGCAATT 480
 QY 481 GGAACAGGAACAGTTCATGCTTCGAGAGAGATGATCCGGAACAGGAGCTAGAAAAAG 540
 DB 481 GGAACAGGAACAGTTCATGCTTCGAGAGAGATGATCCGGAACAGGAGCTAGAAAAAG 540
 QY 541 GCGACTGAAATTTGTAAGAGATTGGGAAGTGAAGCCCTGAGCTAGTGGCCCGCTAGT 600
 DB 541 GCGACTGAAATTTGTAAGAGATTGGGAAGTGAAGCCCTGAGCTAGTGGCCCGCTAGT 600
 QY 601 GCTGACTTGGAGAAAGCCCTCTTGAATGTTCCCACTTAACAGTCTCATCATACA 660
 DB 601 GCTGACTTGGAGAAAGCCCTCTTGAATGTTCCCACTTAACAGTCTCATCATACA 660
 QY 661 GCCTTCAAGCTGTACACACACTGAAGGCCAGTAAACCACTGTGTGAGCAGGTCTT 720
 DB 661 GCCTTCAAGCTGTACACACACTGAAGGCCAGTAAACCACTGTGTGAGCAGGTCTT 720
 QY 721 GAAACCTGAGCACTGAGCAACTCAAGAAATTTCCCAATCATGATGATTTGGCCATGT 780
 DB 721 GAAACCTGAGCACTGAGCAACTCAAGAAATTTCCCAATCATGATGATTTGGCCATGT 780
 QY 781 GGTGGTGGCTGGGGGGCTGGCCCAAGTTTCTCAGTTAGCCAGTCCCAACTGCCCC 840
 DB 781 GGTGGTGGCTGGGGGGCTGGCCCAAGTTTCTCAGTTAGCCAGTCCCAACTGCCCC 840
 QY 841 GGGAGTGGAGACATGTGGAATTTCTGTGAAAACTGATGAGAAATTTTACATTAC 900
 DB 841 GGGAGTGGAGACATGTGGAATTTCTGTGAAAACTGATGAGAAATTTTACATTAC 900
 QY 901 CCAATGTTTCTATCCCAAGCAAGTGTGGGTCTGATTAATGCAACACAGAGAACGAGA 960
 DB 901 CCAATGTTTCTATCCCAAGCAAGTGTGGGTCTGATTAATGCAACACAGAGAACGAGA 960
 QY 961 AGAATCTTTCTCTATACAGATACAGAGGCTCATCACTGGGCTGATTTCTACTCA 1020
 DB 961 AGAATCTTTCTCTATACAGATACAGAGGCTCATCACTGGGCTGATTTCTACTCA 1020

QY 1021 CCCACACAGACCGGCTTTCTCTCCAGTGTGACCTTACACACTCACTGCTTCTTACAGAT 1080
 DB 1021 CCCACACAGACCGGCTTTCTCTCCAGTGTGACCTTACACACTCACTGCTTCTTACAGAT 1080
 QY 1081 GATGTTGACAGAGTACAGAGCCATTTGCTCCCGCAAGTCCAGGAACCTGATTTCTT 1140
 DB 1081 GATGTTGACAGAGTACAGAGCCATTTGCTCCCGCAAGTCCAGGAACCTGATTTCTT 1140
 QY 1141 TAAACTAACTGACATGAGTACAGAGATTTCTCTCTGTCGCAAGAAAGATTTCAATCC 1200
 DB 1141 TAAACTAACTGACATGAGTACAGAGATTTCTCTCTGTCGCAAGAAAGATTTCAATCC 1200
 QY 1201 AACAGCAAGATTCACCTGCTGTTCTGATGAGTGAAGCAGTGAAGTGTGACAGAGC 1260
 DB 1201 AACAGCAAGATTCACCTGCTGTTCTGATGAGTGAAGCAGTGAAGTGTGACAGAGC 1260
 QY 1261 AGTGAACATCAAGACCTTGCATGAGAGCGTTTGAAGTCAACACCTTCCAGAAACAAA 1320
 DB 1261 AGTGAACATCAAGACCTTGCATGAGAGCGTTTGAAGTCAACACCTTCCAGAAACAAA 1320
 QY 1321 CCATATCACTGATCTGTAGCCCTTTAATTTAAGCTTTCTAGAAGCTTTGGAAGTTT 1380
 DB 1321 CCATATCACTGATCTGTAGCCCTTTAATTTAAGCTTTCTAGAAGCTTTGGAAGTTT 1380
 QY 1381 TAGATAGTGAAGGGGGGACATCACTGAGAAAGAGCTGATTTGTATTTGAGGTTTGA 1440
 DB 1381 TAGATAGTGAAGGGGGGACATCACTGAGAAAGAGCTGATTTGTATTTGAGGTTTGA 1440
 QY 1441 AAGAAATATCTGACATATTTTATTTAGGCAAGTCAAGAAAGAAACATGCTACCCAAAAG 1500
 DB 1441 AAGAAATATCTGACATATTTTATTTAGGCAAGTCAAGAAAGAAACATGCTACCCAAAAG 1500
 QY 1501 AACTGTAATCTGAGAAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGAAAGAT 1560
 DB 1501 AACTGTAATCTGAGAAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGAAAGAT 1560
 QY 1561 GGTATATGAACCCCATATACCTTCTCTGATTCACCAATTTGATTAACATTTTTC 1620
 DB 1561 GGTATATGAACCCCATATACCTTCTCTGATTCACCAATTTGATTAACATTTTTC 1620
 QY 1621 CTCTGACATACCTTCTAATTTCTCTAATTTTCAATTTGATTTTATTTTACCTGAGGCT 1680
 DB 1621 CTCTGACATACCTTCTAATTTCTCTAATTTTCAATTTGATTTTATTTTACCTGAGGCT 1680
 QY 1681 CAATAAGGCACTGTGCAAGAAATTTGGAAGCCATTTGAAATCTTTTGAATTTCTG 1740
 DB 1681 CAATAAGGCACTGTGCAAGAAATTTGGAAGCCATTTGAAATCTTTTGAATTTCTG 1740
 QY 1741 TGGTTTATGGAATATGAAATGAGCTTATTAATGAGGAGAGAGTCTTCCATTTG 1800
 DB 1741 TGGTTTATGGAATATGAAATGAGCTTATTAATGAGGAGAGAGTCTTCCATTTG 1800
 QY 1801 ACCAGATGTTGGCTAACAACATCCGGAAGATGATTTTGTACAGAAATTTATTTATTA 1860
 DB 1801 ACCAGATGTTGGCTAACAACATCCGGAAGATGATTTTGTACAGAAATTTATTTATTA 1860
 QY 1861 ATTAATATTTTCAAGATATTTTCTCTCAATTAAGTAAAGTAAATTAACCTTA 1910
 DB 1861 ATTAATATTTTCAAGATATTTTCTCTCAATTAAGTAAAGTAAATTAACCTTA 1910

RESULT 2
 AAF57478
 ID AAF57478 standard; DNA; 2003 BP.
 XX
 AC AAF57478;
 XX
 DT 11-JUN-2001 (first entry)
 XX
 DE Human SAP1/AMSH nucleotide sequence.
 XX
 KM Smad associating polypeptide; SAP; cell cycle; TGF-beta; SAP1/AMSH;

KW signal transduction; phosphorylation; gene therapy; human; ds.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 XX CDS 63..1337
 XX FT /*tag=a
 XX FT /product="SAP1/AMSH"
 XX
 XX MO200121794-A2.
 XX
 XX 29-MAR-2001.
 XX
 XX 20-SEP-2000; 2000WO-US025790.
 XX
 XX 20-SEP-1999; 99US-0154846P.
 XX
 XX (LUDWIG INST CANCER RES.
 XX
 XX Itoh F, Itoh S, Heldin C, Ten-Dijke P;
 XX WPI; 2001-257987/26.
 XX P-PSDB; AAB62193.
 XX
 XX Novel Smad associating proteins useful for regulating cell cycle in a
 XX mammalian cell, and for identifying lead compounds for pharmacological
 XX agents.
 XX
 XX Example 1; Page 53-55; 90pp; English.
 XX
 XX The invention relates to Smad associating polypeptides (SAP) and nucleic
 XX acids encoding the SAPs. The SAPs can be expressed by standard
 XX recombinant methodology. The SAP proteins and polynucleotides are useful
 XX for regulating the cell cycle in a mammalian cell, and for modulating TGF-
 XX beta superfamily signal transduction in a mammalian cell. SAP
 XX polypeptide are useful for testing Smad binding, for testing
 XX phosphorylation, for generating antibodies, and as components of an
 XX immunosassay. The SAP polynucleotides are useful in gene therapy. The
 XX present sequence represents the nucleotide sequence of human SAP1/AMSH
 XX that can be used in modulating TGF-beta superfamily signal transduction
 XX in a mammalian cell
 XX
 XX Sequence 2003 BP; 626 A; 439 C; 441 G; 497 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 1910; DB 5; Length 2003;
 XX Best Local Similarity 100.0%; Pred. No. 0;
 XX Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 CTTGGTCTGATGTCTGACATGAGATGTGAGCTTCCGCCCGAGACCGGGTGAAGGC 60
 XX 53 CTTGGTCTGATGTCTGACATGAGATGTGAGCTTCCGCCCGAGACCGGGTGAAGGC 112
 XX
 XX 61 TCTCTCCAGCTGGGTAGTGGCGTAGAGTGAAGTGAAGATTCACCCCGTGGTACTT 120
 XX 113 TCTCTCCAGCTGGGTAGTGGCGTAGAGTGAAGTGAAGATTCACCCCGTGGTACTT 172
 XX
 XX 121 CCGCTCGAGATTGAGATTATCCGAATGCGATTCATTACTGTGAGAAAGCAATTGA 180
 XX 173 CCGCTCGAGATTGAGATTATCCGAATGCGATTCATTACTGTGAGAAAGCAATTGA 232
 XX
 XX 181 ACATGCTTCATCCTCTATATCAAGTATATCAGCGCTTTTATTGAGAAATTTACAAACA 240
 XX 233 ACATGCTTCATCCTCTATATCAAGTATATCAGCGCTTTTATTGAGAAATTTACAAACA 292
 XX
 XX 241 TCGAGATTACAAATCTGCTGATTCCTGAAAAGAAACACAGTAAGAAATTTAAGA 300
 XX 293 TCGAGATTACAAATCTGCTGATTCCTGAAAAGAAACACAGTAAGAAATTTAAGA 352
 XX
 XX 301 GATTGCAATTTCCCAAGCAGAGAGCTGAAGCGAGCTTTTAAAGATATCCAAAGA 360
 XX 353 GATTGCAATTTCCCAAGCAGAGAGCTGAAGCGAGCTTTTAAAGATATCCAAAGA 412
 XX
 XX 361 ATATACAGATATATATGAAGAAAGAAAGAGAGAGAGATTTGGCCCGAATGCG 420

Db
 413 ATATACAGATATATATGAAGAAAGAAAGAGAGAGAGATTTGGCCCGAATGCG 472
 QY
 421 CATCCAGCAGAGCTGGAAAAGAAAAACAGAGGTATGACACACAGAGAGACAAATT 480
 Db
 473 CATCCAGCAGAGCTGGAAAAGAAAAACAGAGGTATGACACACAGAGAGACAAATT 532
 QY
 481 GAAACAGAAACAGTTCCATGCTTCGAGAGATGATCCGAAACAGAGCTAGAAAAAG 540
 Db
 533 GAAACAGAAACAGTTCCATGCTTCGAGAGATGATCCGAAACAGAGCTAGAAAAAG 592
 QY
 541 GCGACTGAAAATTTGTACAGAGAGTTTGGAAAGTGAACCTTGCTTGAGCCCTAGT 600
 Db
 593 GCGACTGAAAATTTGTACAGAGAGTTTGGAAAGTGAACCTTGCTTGAGCCCTAGT 652
 QY
 601 GCCTGACTTGGAGAAACCTCTCTAGATGTGTCCCACTTAAACAGTCCATCCATACA 660
 Db
 653 GCCTGACTTGGAGAAACCTCTCTAGATGTGTCCCACTTAAACAGTCCATCCATACA 712
 QY
 661 GCCTTCAAGCTGTACACAACTGTAAAGCCAGCTAAAGCACTTGTTGAGACAGTCTT 720
 Db
 713 GCCTTCAAGCTGTACACAACTGTAAAGCCAGCTAAAGCACTTGTTGAGACAGTCTT 772
 QY
 721 GAAACCTGAGAGACTGAGAACTCAGAAAGTATCCCAATTCGATGATGCGCATGT 780
 Db
 773 GAAACCTGAGAGACTGAGAACTCAGAAAGTATCCCAATTCGATGATGCGCATGT 832
 QY
 781 GGTGTGCTGTGGCGGCTGTGCCACAGTTTCTCAGTTAGCCAGTGCACACACTGCCG 840
 Db
 833 GGTGTGCTGTGGCGGCTGTGCCACAGTTTCTCAGTTAGCCAGTGCACACACTGCCG 892
 QY
 841 GGAAGTGGAGACATGTGGAATTTCTGTGTGAAAACATGATGAGAAATTTTACATTTC 900
 Db
 893 GGAAGTGGAGACATGTGGAATTTCTGTGTGAAAACATGATGAGAAATTTTACATTTC 952
 QY
 901 CCATGTTCTCATCCCAACCAAGTGTGGTGTGTTACTGCAACACAGAGAGCAAGAA 960
 Db
 953 CCATGTTCTCATCCCAACCAAGTGTGGTGTGTTACTGCAACACAGAGAGCAAGAA 1012
 QY
 961 AGAATCTTTCCTCATACAGATTAAGAGAGGCTCATACACTGGGCTGATTCATCTCA 1020
 Db
 1013 AGAATCTTTCCTCATACAGATTAAGAGAGGCTCATACACTGGGCTGATTCATCTCA 1072
 QY
 1021 CCCACACAGACCGGCTTCTCTCAGTGTGACCTTACACATCCTGCTTACACAGT 1080
 Db
 1073 CCCACACAGACCGGCTTCTCTCAGTGTGACCTTACACATCCTGCTTACACAGT 1132
 QY
 1081 GATGTTGCCAGAGTGAATGACATTTGTTGCTCCCAAGTTCCAGAACTGATTTCTT 1140
 Db
 1133 GATGTTGCCAGAGTGAATGACATTTGTTGCTCCCAAGTTCCAGAACTGATTTCTT 1192
 QY
 1141 TAAATCTTAATGACATGTGATAGAGAGATTTTCTCTGTGCGCAAGAAAGATTTCTATCC 1200
 Db
 1193 TAAATCTTAATGACATGTGATAGAGAGATTTTCTCTGTGCGCAAGAAAGATTTCTATCC 1252
 QY
 1201 ACACAGCAGAGATTCACCTCTGTTCTGTAGCTGCAGCCACGAGTGTGTTGAGCAGAGC 1260
 Db
 1253 ACACAGCAGAGATTCACCTCTGTTCTGTAGCTGCAGCCACGAGTGTGTTGAGCAGAGC 1312
 QY
 1261 AGTGACCATCACAGACTTCGATGAGCGTTGATGTCAAACACTTTCAGAAACAAATAA 1320
 Db
 1313 AGTGACCATCACAGACTTCGATGAGCGTTGATGATGTCAAACACTTTCAGAAACAAATAA 1372
 QY
 1321 CCATATCAGTGTACGTAGCCCTTAATTTAGCTTTAGAGAGCTTTGGAAGTTTGG 1380
 Db
 1373 CCATATCAGTGTACGTAGCCCTTAATTTAGCTTTTATGAACTTTGGAAGTTTGG 1432
 QY
 1381 TAGATATGTAAGAAAGGGGGGCACTACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGA 1440
 Db
 1433 TAGATATGTAAGAAAGGGGGGCACTACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGA 1492
 QY
 1441 AAGAAATTAAGTGAATATTTTATAGGCAAGTCAGAAAGAGATGCTACCCCAAGC 1500

Db 901 AGCTTATTACTGGGAGGAGACGCTTACCTCATTGACGAGATTTGTTGGCTTAACACA 960
Qy 1823 TCCCGAAGATGATTTTGTCAAGATTTATTTATTAATAATTTCAAGATATTTT 1882
Db 961 TCCCGAAGATGATTTTGTCAAGATTTATTTATTAATAATTTCAAGATATTTT 1020
Qy 1883 CCTCACAATAAAGTAACAATTA 1905
Db 1021 CCTCACAATAAAGTAACAATTA 1043

RESULT 4
ABN72603
ID ABN72603 standard; DNA; 1043 BP.
XX
AC ABN72603;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ovarian carcinoma antigen polynucleotide #19.
XX
KM Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US022635.
XX
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vealwick TS, Carter D, Hill F, Abdone E;
XX
DR WPI; 2002-164781/21.
XX
XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
PS Example 1; Page 99; 408bp; English.
XX
XX This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytosolic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents DNA related
CC to the invention
XX
SQ Sequence 1043 BP; 313 A; 224 C; 198 G; 308 T; 0 U; 0 Other;

Query Match 54.6%; Score 1043; DB 6; Length 1043;
Best Local Similarity 100.0%; Pzed. No. 8.8e-276;
Matches 1043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 863 CTCTGTGGAACATGATGAGAAATGATTTACATTACCATGCTTCATCCCAAGCAA 922
Db 1 CTCTGTGGAACATGATGAGAAATGATTTACATTACCATGCTTCATCCCAAGCAA 60
Qy 923 AGTGTGGGTCTGATTACTGCAACACAGAGAAGAAAGAACTTTTCTCATACAGAT 982
Db 61 AGTGTGGGTCTGATTACTGCAACACAGAGAAGAAAGAACTTTTCTCATACAGAT 120

Qy 983 CAGCAGGGCTCATCACTGGGCTGGAATTCATCTACCCCAACAGACCGGTTCTC 1042
Db 121 CAGCAGGGCTCATCACTGGGCTGGAATTCATCTACCCCAACAGACCGGTTCTC 180
Qy 1043 TCCAGTGTGACCTTACACACTGCTCTTACCAATATGTTGGCAGAGTAGAGCC 1102
Db 181 TCCAGTGTGACCTTACACACTGCTCTTACCAATATGTTGGCAGAGTAGAGCC 240
Qy 1103 ATTGTTGCTCCCGCAAGTTCAGAAACTGATTCCTTAACTAAGTACCATGACTA 1162
Db 241 ATTGTTGCTCCCGCAAGTTCAGAAACTGATTCCTTAACTAAGTACCATGACTA 300
Qy 1163 GAGAGATTTCTTCTCTGTGCGCCAGAAAGATTTTCATCCACACAGAGATTCACCTTG 1222
Db 301 GAGAGATTTCTTCTCTGTGCGCCAGAAAGATTTTCATCCACACAGAGATTCACCTTG 360
Qy 1223 TTCTGTAGCTGACGACGCTGACTGTTTGTGACAGAGAGTGCATCAGACCTTGA 1282
Db 361 TTCTGTAGCTGACGACGCTGACTGTTTGTGACAGAGAGTGCATCAGACCTTGA 420
Qy 1283 TGAAGCTTTGAGTCCAAACCTTCCAAAGAACAAACCATATCATGTTAGTGGCC 1342
Db 421 TGAAGCTTTGAGTCCAAACCTTCCAAAGAACAAACCATATCATGTTAGTGGCC 480
Qy 1343 CTTAATTTAGCTTCTAGAAAGCTTTTGAAGTTTGTAGATAGTAAAGGGGGGCT 1402
Db 481 CTTAATTTAGCTTCTAGAAAGCTTTTGAAGTTTGTAGATAGTAAAGGGGGGCT 540
Qy 1403 CACCTGAGAAAGAGCTGATTTTGTATTCAGGTTTGAAGAAATACGAAATATTTT 1462
Db 541 CACCTGAGAAAGAGCTGATTTTGTATTCAGGTTTGAAGAAATACGAAATATTTT 600
Qy 1463 TTAGGCAAGTCAAGAAAGAAATATGTCACCCCAAGCACTGTAACTCAGAAATTAAGT 1522
Db 601 TTAGGCAAGTCAAGAAAGAAATATGTCACCCCAAGCACTGTAACTCAGAAATTAAGT 660
Qy 1523 TACTCAGAAATTAAGTGTGCTCAGAAATTAAGAAAGATGTTAATGAACCCCATATAC 1582
Db 661 TACTCAGAAATTAAGTGTGCTCAGAAATTAAGAAAGATGTTAATGAACCCCATATAC 720
Qy 1583 CCTTCCTTGTGATTCACCAATGTTTAAATTTTTCCTCAGCTATCCTTAAATTT 1642
Db 721 CCTTCCTTGTGATTCACCAATGTTTAAATTTTTCCTCAGCTATCCTTAAATTT 780
Qy 1643 CTCTCAATTTCAATTTGTTTATTTACCTGCGGCTCAATAAGGATCTGTCCAGAA 1702
Db 781 CTCTCAATTTCAATTTGTTTATTTACCTGCGGCTCAATAAGGATCTGTGTCCAGAA 840
Qy 1703 ATTGGAAAGCCATTTAGAAATCTTTTGATTTTCTGTGTTTATGGCAATATGATGG 1762
Db 841 ATTGGAAAGCCATTTAGAAATCTTTTGATTTTCTGTGTTTATGGCAATATGATGG 900
Qy 1763 AACTTATTTACTGGGTGAGGAGACACTTATTCATTTGACAGATTTTGGCTAACCA 1822
Db 901 AACTTATTTACTGGGTGAGGAGACACTTATTCATTTGACAGATTTTGGCTAACCA 960
Qy 1823 TCCCGAAGATGATTTTGTCAAGATTTATTTATTAATAATTTCAAGATATTTT 1882
Db 961 TCCCGAAGATGATTTTGTCAAGATTTATTTATTAATAATTTCAAGATATTTT 1020

RESULT 5
ADA08768
ID ADA08768 standard; DNA; 1043 BP.
XX
AC ADA08768;
XX
DT 06-NOV-2003 (first entry)
XX

DE Human ovarian carcinoma antigen polynucleotide #19.
 XX de; human; gene therapy; ovarian cancer; cancer.
 KM Homo sapiens.
 XX US2003091580-A1.
 XX 15-MAY-2003.
 XX 17-JUL-2001; 2001US-00907969.
 XX 18-JUN-2001; 2001US-00884441.
 XX (MITC/) MITCHAM J L.
 XX (KING/) KING G E.
 XX (ALGA/) ALGATE P A.
 XX (FLIN/) FLING S P.
 XX (RETT/) RETTER M W.
 XX (FANG/) FANGER G R.
 XX (REED/) REED S G.
 XX (VEDV/) VEDVICK T S.
 XX (CART/) CARTER D.
 XX (HILL/) HILL P.
 XX (ALBO/) ALBONE E.
 XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
 FI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
 XX WPI; 2003-532352/50.
 XX New isolated O72P polypeptides and polynucleotides, useful in gene
 PT therapy, particularly for treating or diagnosing cancer, e.g. ovarian
 PT cancer.
 XX Example 1; Fig 1; 371p; English.
 XX The invention relates to an isolated O72P polypeptide, which has the
 CC structure fully defined in the specification. The composition containing
 CC the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
 CC or antigen presenting cells are useful for stimulating an immune response
 CC and treating ovarian cancer. Detecting the presence of the
 CC polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
 CC carcinoma cDNAs and protein cDNAs were identified using microarray
 CC technology. The present sequence represents a human ovarian carcinoma
 CC antigen polynucleotide.
 XX Sequence 1043 BP; 313 A; 224 C; 198 G; 308 T; 0 U; 0 Other;
 SQ
 Query Match 54.6%; Score 1043; DB 8; Length 1043;
 Best Local Similarity 100.0%; Pred. No. 8.8e-276;
 Matches 1043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1163 GAGGAGATTTCTCTGCGCAGAAAGATTTCATCCACACAGCAAGATCCACTCTG 1222
 DB 301 GAGAGATTCTTCTGTCGCCAGAAAGATTTCATCCACACAGCAAGATCCACTCTG 360
 QY 1223 TTCTGAGCTGACACGACGTAAGTTGTGAGACAGAGAGTACATCAAGACCTTGA 1282
 DB 361 TTCTGAGCTGACACGACGTAAGTTGTGAGAGAGAGTACATCAAGACCTTGA 420
 QY 1283 TGAGCGTTGAGTCCACACCTTCCAGAAACAAACAAACATTCAGTACTGTAGCCC 1342
 DB 421 TGAGCGTTGAGTCCACACCTTCCAGAAACAAACAAACATTCAGTACTGTAGCCC 480
 QY 1343 CTTAATTTAAGCTTTGAGAAAGCTTGAAGTTTGTGATGATGAGAAAGGGGGCAT 1402
 DB 481 CTTAATTTAAGCTTTGAGAAAGCTTGAAGTTTGTGATGATGAGAAAGGGGGCAT 540
 QY 1403 CACCTGAGAAAGAGCTGATTTTGTATTCAGTTTGAAGAAATTAAGTACATATTT 1462
 DB 541 CACCTGAGAAAGAGCTGATTTTGTATTCAGTTTGAAGAAATTAAGTACATATTT 600
 QY 1463 TTAGGCAAGTCAGAAAGAGACATGCTCACCAGAAACAACTGTACTAGAAATTAAGT 1522
 DB 601 TTAGGCAAGTCAGAAAGAGACATGCTCACCAGAAACAACTGTACTAGAAATTAAGT 660
 QY 1523 TACTCAGAAATTAAGTACGCTCAGAAATTAAGAAAGAAATGATTAAGACCCCATATAC 1582
 DB 661 TACTCAGAAATTAAGTACGCTCAGAAATTAAGAAAGAAATGATTAAGACCCCATATAC 720
 QY 1583 CCTTCCTGAGATTCACCAATGTTTATTAATTTTCTGCTCAGTATCCTTTAAATTT 1642
 DB 721 CCTTCCTGAGATTCACCAATGTTTATTAATTTTCTGCTCAGTATCCTTTAAATTT 780
 QY 1643 CTCTCTAATTTCAATTTGTTTATTAATTTAATTTCTGCGCTCAATTAAGGCAATGTCAGAA 1702
 DB 781 CTCTCTAATTTCAATTTGTTTATTAATTTAATTTCTGCGCTCAATTAAGGCAATGTCAGAA 840
 QY 1703 ATTGGAGCAATTTAGAAATCTTTGGATTTTCTGCTGATTTATGCAATATGAATGG 1762
 DB 841 ATTGGAGCAATTTAGAAATCTTTGGATTTTCTGCTGATTTATGCAATATGAATGG 900
 QY 1763 AGCTTATTACTGCGGTGAGGAGACAGCTTACATTTGACCAAGATGTTGGCTTAACACA 1822
 DB 901 AGCTTATTACTGCGGTGAGGAGACAGCTTACATTTGACCAAGATGTTGGCTTAACACA 960
 QY 1823 TCCCGAAGAAATGATTTTGTGAGAAATTTGTTTAAATTAATTAATTAAGATATTTT 1882
 DB 961 TCCCGAAGAAATGATTTTGTGAGAAATTTGTTTAAATTAATTAATTAAGATATTTT 1020
 QY 1883 CCTTACATATAAGTAACATTA 1905
 DB 1021 CCTTACATATAAGTAACATTA 1043

RESULT 6
 AAA69708
 ID AAA69708 standard; cDNA; 1041 BP.
 XX AAA69708;
 AC 07-NOV-2000 (first entry)
 XX
 DT
 XX
 DB Human ovarian carcinoma antigen polynucleotide SEQ ID NO:18.
 XX Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
 KM tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
 XX Homo sapiens.
 OS
 XX
 PN WO200036107-A2.
 XX
 PD 22-JUN-2000.
 XX
 PF 17-DEC-1999; 99WO-US030270.


```

XX 17-DEC-1998; 98US-00215661.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00318933.
XX 24-SEP-1999; 99US-00404879.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Frudakis TW;
XX WPI, 2000-431589/37.
XX
PT Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT encoding it, useful for the diagnosis, prevention and treatment of
XX cancer, preferably ovarian cancer.
XX
PS Claim 1; Fig 1; 29pp; English.
XX
CC The present invention describes an isolated polypeptide comprising an
CC immunogenic portion of an ovarian carcinoma protein (or its variants).
CC Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC cancer. AA69691 to AA70077 and AB12552 to AB12557 represent human
CC ovarian carcinoma polynucleotides and proteins used in the
CC exemplification of the present invention
XX
SQ Sequence 1041 BP, 312 A; 223 C; 198 G; 307 T; 0 U; 1 Other;
XX
Query Match 54.5%; Score 1040; DB 3; Length 1041;
Best Local Similarity 99.9%; Pred. No. 5.9e-275;
Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 863 CTCGTGGAAGAGTGAAGATGAATTTACATTACCCATGTTCTCATCCCAAGCAA 922
Db 1 CTCGTGGAAGAGTGAAGATGAATTTACATTACCCATGTTCTCATCCCAAGCAA 60
QY 923 AGTGGCTGGTCTGATTTCTGCAACAGAGAAACCAAAACCTTTTCTCATACAGAT 962
Db 61 AGTGGCTGGTCTGATTTCTGCAACAGAGAAACCAAAACCTTTTCTCATACAGAT 120
QY 983 CAGCAGGAGCTCATACACTGGGCTGGATTCAATCACTACCCCAACAGACCGGATTTCTC 1042
Db 121 CAGCAGGAGCTCATACACTGGGCTGGATTCAATCACTACCCCAACAGACCGGATTTCTC 180
QY 1043 TCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTTGCCAGAGTAGAGCC 1102
Db 181 TCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTTGCCAGAGTAGAGCC 240
QY 1103 ATTGTTGCTCCCAAGTTCAGGAAATGATGATCTTTAACTAATCTGACATGAGCTA 1162
Db 241 ATTGTTGCTCCCAAGTTCAGGAAATGATGATCTTTAACTAATCTGACATGAGCTA 300
QY 1163 GAGGAGATTTCTTCTGTCGCCAGAAAGATTTCAATCCACACAGCAAGATCCACTCTG 1222
Db 301 GAGGAGATTTCTTCTGTCGCCAGAAAGATTTCAATCCACACAGCAAGATCCACTCTG 360
QY 1223 TTCTGTAGCTGACCAAGCTGATGTTGTGAGACAGAGCAAGTGCATCAGACCTTGA 1282
Db 361 TTCTGTAGCTGACCAAGCTGATGTTGTGAGACAGAGCAAGTGCATCAGACCTTGA 420
QY 1283 TGAGCGTTTGAGTCCAAACCTTCCAGAAACAACAAACCATATCAGTATCTGAGCC 1342
Db 421 TGAGCGTTTGAGTCCAAACCTTCCAGAAACAACAAACCATATCAGTATCTGAGCC 480
QY 1343 CTTAATTTAAGCTTTCTGAGAAAGCTTTGAGATTTTGTAGATGAGAAAGGGGGGCA 1402
Db 481 CTTAATTTAAGCTTTCTGAGAAAGCTTTGAGATTTTGTAGATGAGAAAGGGGGGCA 540
QY 1403 CACCTGAGAAAGAGCTGATTTGTATTTGAGTTGAGAAAGAAATACATGAAATATTT 1462
Db 541 CACNTGAGAAAGAGCTGATTTGTATTTGAGTTGAGAAAGAAATACATGAAATATTT 600

```

```

QY 1463 TTAGGCAAGTCAGAAAGAGACATGTCACCCAAAGCAACTGTAACTAGAAATTAAGT 1522
Db 601 TTAGGCAAGTCAGAAAGAGACATGTCACCCAAAGCAACTGTAACTAGAAATTAAGT 660
QY 1523 TACTCAGAAATTAAGTAGCTCAGAAATTAAGAAAGATGATATGAAACCCCATATAC 1582
Db 661 TACTCAGAAATTAAGTAGCTCAGAAATTAAGAAAGATGATATGAAACCCCATATAC 720
QY 1583 CCTTCCTTCGAGTTACCAATTTGTAACAATTTTTCCTTCGACATATCTTCTAATTT 1642
Db 721 CCTTCCTTCGAGTTACCAATTTGTAACAATTTTTCCTTCGACATATCTTCTAATTT 780
QY 1643 CTCTTAATTTCAATTTGTTATATTTACCTCTGAGCTCAATTAAGGAGCATGTGAGAA 1702
Db 781 CTCTTAATTTCAATTTGTTATATTTACCTCTGAGCTCAATTAAGGAGCATGTGAGAA 840
QY 1703 ATTTGAGAGCATTTAGAAATCTTTTGGATTTTCTGTGTTTATGCAATATGATGG 1762
Db 841 ATTTGAGAGCATTTAGAAATCTTTTGGATTTTCTGTGTTTATGCAATATGATGG 900
QY 1763 AGCTTAATTCGAGGAGAGGACAGCTTACCTCATTTGSCAGATTTGCTTAACACA 1822
Db 901 AGCTTAATTCGAGGAGAGGACAGCTTACCTCATTTGSCAGATTTGCTTAACACA 960
QY 1823 TCCCGAAGATGATTTTGTACAGAAATTAATTTAATAATATTTTCAGATATTTT 1882
Db 961 TCCCGAAGATGATTTTGTACAGAAATTAATTTAATAATATTTTCAGATATTTT 1020
QY 1883 CCTCTAATTAAGTAACAT 1903
Db 1021 CCTCTAATTAAGTAACAT 1041
XX
RESULT 7
ABN72602
ID ABN72602 standard; DNA; 1041 BP.
XX
AC ABN72602;
XX
DT 02-JUL-2002 (first entry)
XX
DE Ovarian carcinoma antigen polynucleotide #18.
XX
KW Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US022635.
XX
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, King GE, Algate PA, Rling SP, Retter MW, Fanger GR;
PI Reed SG, Vedrick TS, Carter D, Hill P, Albone E;
XX WPI, 2002-164781/21.
XX
PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT protein or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
PS Example 1; Page 98; 408pp; English.
XX

```

CC This invention relates to polypeptides comprising an immunogenic portion
 CC of an ovarian carcinoma protein which acts as an immunostimulant and is
 CC cytotoxic. The polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations and antigen presenting cells that express
 CC the polypeptides are useful for stimulating an immune response in a
 CC patient and treating ovarian cancer. This sequence represents DNA related
 CC to the invention.

XX
 SQ Sequence 1041 BP; 312 A; 223 C; 198 G; 307 T; 0 U; 1 Other;

Query Match 54.5%; Score 1040; DB 6; Length 1041;

Best Local Similarity 99.9%; Pred. No. 5.9e-275; Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 863 CTCTGTGAAAACGTGATGAGAAATTTTACCACTTACCACTTCTCATCCCAAGCAA 922
DB 1 CTCTGTGAAAACGTGATGAGAAATTTTACCACTTACCACTTCTCATCCCAAGCAA 60
QY 923 AATGCTGGGCTGATTTACTGCAACACAGAAACGAAAGAACTTTCTCATCAGAT 982
DB 61 AATGCTGGGCTGATTTACTGCAACACAGAAACGAAAGAACTTTCTCATCAGAT 120
QY 983 CAGCAGAGGCTCATCATCACTGGGCTGATTCATCACTCAACCAAGACCGCTTTCTC 1042
DB 121 CAGCAGAGGCTCATCATCACTGGGCTGATTCATCACTCAACCAAGACCGCTTTCTC 180
QY 1043 TCCAGTGTGACCTACACACTCACTGCTCTTACCAAGATGTTGCCAGATCAGTAGCC 1102
DB 181 TCCAGTGTGACCTACACACTCACTGCTCTTACCAAGATGTTGCCAGATCAGTAGCC 240
QY 1103 ATTTGTTGCTCCCAAGTTCAGAGAACTGGATTTCTTAACTAACCTGACATGAGACTA 1162
DB 241 ATTTGTTGCTCCCAAGTTCAGAGAACTGGATTTCTTAACTAACCTGACATGAGACTA 300
QY 1163 GAGGAAATTTCTCCGTCGACAGAAAGATTTTACCAACACAGAGATCCACCTCTCG 1222
DB 301 GAGGAAATTTCTCCGTCGACAGAAAGATTTTACCAACACAGAGATCCACCTCTCG 360
QY 1223 TTCTGTAGCTGACGACGACGTCGCTGTTGTGACAGAGACGATCAAGACCTTCTGA 1282
DB 361 TTCTGTAGCTGACGACGACGTCGCTGTTGTGACAGAGACGATCAAGACCTTCTGA 420
QY 1283 TGAGGCTTGTGATCCAAACCTTCCAAACAAACCAATACCTGATGATGATGATGCC 1342
DB 421 TGAGGCTTGTGATCCAAACCTTCCAAACAAACCAATACCTGATGATGATGATGCC 480
QY 1343 CTTAATTTAAGCTTTCTAGAAAGCTTTTGTAGATAGTAAAGGAGGAGCAT 1402
DB 481 CTTAATTTAAGCTTTCTAGAAAGCTTTTGTAGATAGTAAAGGAGGAGCAT 540
QY 1403 CACCTGAGAAAGAGCTGATTTTGTATTCAGATTTGAAAAGAAATTAACATATTTT 1462
DB 541 CACCTGAGAAAGAGCTGATTTTGTATTCAGATTTGAAAAGAAATTAACATATTTT 600
QY 1463 TTAGGCAAGTCAGAAAGAGACATGTCACCAAGAAAGCACTGTACTCAGAAATTAAGT 1522
DB 601 TTAGGCAAGTCAGAAAGAGACATGTCACCAAGAAAGCACTGTACTCAGAAATTAAGT 660
QY 1523 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGATATATGAAACCCCATATAC 1582
DB 661 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGATATATGAAACCCCATATAC 720
QY 1583 CCTCTCTCTGATTCACCAATTTGTTAATTTTCTCTCAGTATCTTCTATATTT 1642
DB 721 CCTCTCTCTGATTCACCAATTTGTTAATTTTCTCTCAGTATCTTCTATATTT 780
QY 1643 CTCTCTAATTTCAATTTGTTATATTTCCCTCTGGGCTCAATTAAGGACATCTGTGAGAA 1702
DB 781 CTCTCTAATTTCAATTTGTTATATTTACCTCTGGGCTCAATTAAGGACATCTGTGAGAA 840
QY 1703 ATTGGAAGCCATTAGAAAATCTTTGATTTCTGTGGTTATGCAATATGATG 1762
DB 841 ATTGGAAGCCATTAGAAAATCTTTGATTTCTGTGGTTATGCAATATGATG 900

```

```

QY 1763 AGCTTATTTAGGGGTGAGGACAGCTTACCTCATTTGACACAGATTTGGCTAACACA 1822
DB 901 AGCTTATTTAGGGGTGAGGACAGCTTACCTCATTTGACACAGATTTGGCTAACACA 960
QY 1823 TCCCGAAGATGATTTTGTGCGAATTTGTTATTTAATTAATATTTTCAGATATTTT 1882
DB 961 TCCCGAAGATGATTTTGTGCGAATTTGTTATTTAATTAATATTTTCAGATATTTT 1020
QY 1883 CCTTACATTAAGTAACTAAT 1903
DB 1021 CCTTACATTAAGTAACTAAT 1041

```

RESULT 8

ADA08767
 ID ADA08767 standard; DNA; 1041 BP.

XX AC ADA08767;

XX DT 06-NOV-2003 (first entry)

XX DE Human ovarian carcinoma antigen polynucleotide #18.

XX KW de; human; gene therapy; ovarian cancer; cancer.

XX OS Homo sapiens.

XX EN US2003091580-A1.

XX PD 15-MAY-2003.

XX PF 17-JUL-2001; 2001US-00907969.

XX PR 18-JUN-2001; 2001US-00884441.

XX PA (MTC)/ MITCHAM J L.

XX PA (KING)/ KING G E.

XX PA (ALGA)/ ALGATE P A.

XX PA (FLIN)/ FLING S P.

XX PA (RETT)/ RETTER M W.

XX PA (FANG)/ FANGER G R.

XX PA (REED)/ REED S G.

XX PA (VEDV)/ VEDVICK T S.

XX PA (CART)/ CARTER D.

XX PA (HILL)/ HILL P.

XX PA (ALBO)/ ALBONE E.

XX PI Micham J L, King G E, Algate P A, Retter M W, Fanger G R;

XX PT Reed S G, Vedvick T S, Carter D, Hill P, Albone E;

XX PS WPI; 2003-532352/50.

XX DR WPI; 2003-532352/50.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

Query Match 54.5%; Score 1040; DB 8; Length 1041;
 Best Local Similarity 99.9%; Pred. No. 5.9e-275;

Example 1; Fig 1; 371bp; English.
 The invention relates to an isolated 0772p polypeptide, which has the structure fully defined in the specification. The composition containing the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells or antigen presenting cells are useful for stimulating an immune response and treating ovarian cancer. Detecting the presence of the polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian carcinoma cDNAs and protein cDNAs were identified using microarray technology. The present sequence represents a human ovarian carcinoma antigen polynucleotide.

Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 863 CTCGTGTGAAAACTGATGAGAAATGATTTTACATTAACCATTCATGCTTCATCCCAAGCAA 922
DB 1 CTCGTGTGAAAACTGATGAGAAATGATTTTACATTAACCATTCATGCTTCATCCCAAGCAA 60
QY 923 AGTGTGAGTCTGATGATGACACACAGAAACCAAGAGAACTTTCCATCATCAGGAT 982
DB 61 AGTGTGAGTCTGATGATGACACACAGAAACCAAGAGAACTTTCCATCATCAGGAT 120
QY 983 CAGCAGGAGCTTCATCAGACTGGGCTGATGATTAATCACTACCCACAGACAGCCGTTTCTC 1042
DB 121 CAGCAGGAGCTTCATCAGACTGGGCTGATGATTAATCACTACCCACAGACAGCCGTTTCTC 180
QY 1043 TCCAGTGTGACACTACACATCACTGCTTACCAAGATGATGTCAGAGTCACTAGCC 1102
DB 181 TCCAGTGTGACACTACACATCACTGCTTACCAAGATGATGTCAGAGTCACTAGCC 240
QY 1103 ATTGTTTGTCTCCCAAGTTCAGAGAACTGATTTCTTTAACTAACTAGCACTAGACTA 1162
DB 241 ATTGTTTGTCTCCCAAGTTCAGAGAACTGATTTCTTTAACTAACTAGCACTAGACTA 300
QY 1163 GAGGAGATTTCTCTCTGTCGCCAAGAAAGATTTTATCCACACAGAGATCCACTCTG 1222
DB 301 GAGGAGATTTCTCTCTGTCGCCAAGAAAGATTTTATCCACACAGAGATCCACTCTG 360
QY 1223 TTCGTAGCTGAGCAGCAGTGACTGTTGTGACAGAGCAGTGAACCATCAAGACCTTGA 1282
DB 361 TTCGTAGCTGAGCAGCAGTGACTGTTGTGACAGAGCAGTGAACCATCAAGACCTTGA 420
QY 1283 TGAGGCTTTGAGTCCACACACTTCCAGAACACAAACCATATCATGATGATCTAGCCC 1342
DB 421 TGAGGCTTTGAGTCCACACACTTCCAGAACACAAACCATATCATGATGATCTAGCCC 480
QY 1343 CTAAATTTAGCTTCTGAGAAAGCTTTGAGAGTTTGTATATGTGAGAAAGGCGGCGAT 1402
DB 481 CTAAATTTAGCTTCTGAGAAAGCTTTGAGAGTTTGTATATGTGAGAAAGGCGGCGAT 540
QY 1403 CACCTGAGAAAGCTGATTTTGTATTTTCAGGTTTGAAGAAATTAATCTGAACATATTT 1462
DB 541 CACCTGAGAAAGCTGATTTTGTATTTTCAGGTTTGAAGAAATTAATCTGAACATATTT 600
QY 1463 TTAGGCAAGTCAAGAAAGAAATGTCACCCAAAGCACTGTAATCTGAGAAATTAAGT 1522
DB 601 TTAGGCAAGTCAAGAAAGAAATGTCACCCAAAGCACTGTAATCTGAGAAATTAAGT 660
QY 1523 TACTCAGAAATTAAGTACTGAGAAATTAAGAAATGATATATGAACCCCATATAC 1582
DB 661 TACTCAGAAATTAAGTACTGAGAAATTAAGAAATGATATATGAACCCCATATAC 720
QY 1583 CCTTCCTCTGAGTTCACCAATTTGTTAATTTTCTCTGCTGATCCCTTATATTT 1642
DB 721 CCTTCCTCTGAGTTCACCAATTTGTTAATTTTCTCTGCTGATCCCTTATATTT 780
QY 1643 CTCTCTAATTTCAATTTGTTATATTTTACCTCTGGGCTCAATAGGCGATCTGCGAGAA 1702
DB 781 CTCTCTAATTTCAATTTGTTATATTTTACCTCTGGGCTCAATAGGCGATCTGCGAGAA 840
QY 1703 ATTGGAAGCCATTTAGAAATCTTTGATTTTCTCTGCTGTTATAGGCAATAGATGG 1762
DB 841 ATTGGAAGCCATTTAGAAATCTTTGATTTTCTCTGCTGTTATAGGCAATAGATGG 900
QY 1763 AGCTTATTAATGAGGAGTGAAGGACAGCTTACTCAATTTGACCAAGATGTTGGCTAACCA 1822
DB 901 AGCTTATTAATGAGGAGTGAAGGACAGCTTACTCAATTTGACCAAGATGTTGGCTAACCA 960
QY 1823 TCCCGAAGAGATTTTGTGAGAGATTTATGTTATTTATATAATTTGAGATTTT 1882
DB 961 TCCCGAAGAGATTTTGTGAGAGATTTATGTTATTTATATAATTTGAGATTTT 1020
QY 1883 CCTTACATTAATTAAGTACAT 1903
DB 1021 CCTTACATTAATTAAGTACAT 1041

```

```

RESULT 9
AAA59853
ID AAA59853 standard; cDNA, 1384 BP.
XX
AC AAA59853;
XX
DT 13-OCT-2000 (first entry)
XX
DE Murine AMSH nucleotide sequence.
XX
KW AMSH; signal transduction; SH3 domain; cytokine; STAM; severe infection;
KW cancer; autoimmune disorder; mouse; ss.
XX
OS Mus sp.
XX
WO200029436-A1.
XX
PD 25-MAY-2000.
XX
PF 12-NOV-1999; 99MO-JP006309.
XX
PR 12-NOV-1998; 98UP-00322674.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Sugamura K, Tanaka N;
XX
DR WPI: 2000-387743/33.
XX
DR P-PSDB; AAB03792.
XX
PT Protein AMSH being signal transducer interacting with SH3 domain of
PT cytokine signal transducer STAM, for diagnosis and development of drugs
PT of treating diseases e.g. severe infection, cancers and autoimmune
PT disorders.
XX
PS Claim 10; Page 18-19; 22pp; Japanese.
XX
CC This sequence represents the murine AMSH gene. The AMSH protein is
CC involved in signal transduction, and interacts with the SH3 domain of
CC cytokine signal transducer STAM. The invention includes human and murine
CC AMSH protein and gene sequences (hAMSH and mAMSH), and also includes
CC antibodies directed against hAMSH and mAMSH. The AMSH protein can be used
CC in the diagnosis and development of drugs for treating diseases due to
CC the dysfunction of cytokine signal transduction e.g. severe infection,
CC cancers and autoimmune disorders
XX
SQ Sequence 1384 BP; 402 A; 354 C; 323 G; 305 T; 0 U; 0 Other;

Query Match 46.1%; Score 879.8; DB 3; Length 1384;
Best Local Similarity 80.6%; Pred. No. 7.4e-231;
Matches 1088; Conservative 0; Mismatches 247; Indels 0; Gaps 0;
QY 11 ATGTCTGACCATGAGATGTAGGCTTCCGCCGAGAAACCGGCTGAGGCTCTCTCCAG 70
DB 58 ATGTCTGACCATGAGGAGTGAAGCTTCCACCCCAAGACCGGCTGAGGATTTCTGTCCAA 115
QY 71 CTGGGATGAGGCGGTGAGTGAATGAAGACATTCACCCGCTGGTACTCCGCTCGA 130
DB 116 CTGGGAGTGAAGTGAATGAAGACATTCACCCGCTGGTACTCCGCTCGGAT 175
QY 131 GTTGAATTAATCCGATGATGATTCATTTACTCTGAGGAGGCAATGAAATGATGCTTC 190
DB 176 GTTGAATTAATCCGATGATGATTCATTTACTCTGAGGAGGCAATGAAATGATGCTTC 235
QY 191 ATCCCTTAATCAATGATATCAAGCTCTTTATGAGAAACATCAACAAATCGAGTATAC 250
DB 236 ATCCCTTAATCAATGATATCAAGCTCTTTATGAGAAACATTCGAGAAACATCGAGTATAC 295
QY 251 AATCTGCTGATCTCTGAGAAAGAAAGACACAGTAAAGAAATTTAAAGGATTTGATTT 310
DB 296 AATCTGATATCATCTCTGAGAAAGAAAGATGCTGTCAAGAAATTTAAAGGCTGCTTTC 355

```


QY 1589 TTCTGGATTCACCAATTTGTAACATTTTTCCTCTCAGTATCTCTTAATTTCTCT 1648
 Db 361 TTCTGGATTCACCAATTTGTAACATTTTTCCTCTCAGTATCTCTTAATTTCTCT 420
 QY 1649 AATTTCATTTGTTATATTTTAACTCTGCGCTCAATTAAGGCAATCTGCAAGAAATTTGG 1708
 Db 421 AATTTCATTTGTTATATTTTAACTCTGCGCTCAATTAAGGCAATCTGCAAGAAATTTGG 480
 QY 1709 AAGCCATTTAGAAATCTTTGATTTTCTGCTGTTATAGCAATATGATGAGCTTA 1768
 Db 481 AAGCCATTTAGAAATCTTTGATTTTCTGCTGTTATAGCAATATGATGAGCTTA 540
 QY 1769 TTACTGGGCTGAGGAGCAAGCTTACTCTCAATTTGACCAATTTGCTTACACATCCCGA 1828
 Db 541 TTACTGGGCTGAGGAGCAAGCTTACTCTCAATTTGACCAATTTGCTTACACATCCCGA 600
 QY 1829 AGAATGATTTTTCGCGGAATTTTGTATTTTAAATTTTCAAGGATTTTTCCTCTA 1888
 Db 601 AGAATGATTTTTCGCGGAATTTTGTATTTTAAATTTTCAAGGATTTTTCCTCTA 660
 QY 1889 CATAAAGTAACTAATTAATTA 1910
 Db 661 CATAAAGTAACTAATTAATTA 682

RESULT 11
 ID AAS85218 standard; cDNA; 1236 BP.
 AC AAS85218;
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #21022.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 PN MO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US008631.
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HUSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG21031.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 21022; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridization probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX

Sequence 1236 BP; 331 A; 317 C; 338 G; 250 T; 0 U; 0 Other;
 SQ
 Query Match 33.2%; Score 634.4; DB 5; Length 1236;
 Best Local Similarity 99.8%; Pred. No. 1.9e-163;
 Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 380 GAAAAAGAAAGAAAGAGAGAAATTTGCCCCGAAATGCCCCATCCAGACAGAGCTTGA 439
 Db 250 GAGAAAGAAAGAAAGAGAGAAATTTGCCCCGAAATGCCCCATCCAGACAGAGCTTGA 309
 QY 440 AAGGAAAAACAGAGAGTATGACACACAGAGAGAGCAATTTGGACAGAAACGTTCCAT 499
 Db 310 AAGGAAAAACAGAGAGTATGACACACAGAGAGAGCAATTTGGACAGAAACGTTCCAT 369
 QY 500 GCTTTCAGAGATGATTCGGAATCCAGAGACCTTGAAGAAAGAGAGTGAATTTGACG 559
 Db 370 GCTTTCAGAGATGATTCGGAATCCAGAGACCTTGAAGAAAGAGAGTGAATTTGACG 429
 QY 560 GAGTTTGGAGAGTAGACCTGCGCTAGTGGCCGCTAGTGGCTGACTTGGAGAAAGCC 619
 Db 430 GAGTTTGGAGAGTAGACCTGCGCTAGTGGCCGCTAGTGGCTGACTTGGAGAAAGCC 489
 QY 620 TCCTTAGATGTTTCCCACTTAACTGCTCATTCATACAGCTTCAAGCTTCAAGCTTCA 679
 Db 490 TCCTTAGATGTTTCCCACTTAACTGCTCATTCATACAGCTTCAAGCTTCAAGCTTCA 549
 QY 680 ACTGTAAGGCGAGGTAAGACCACTGCTGAGAGAGGCTTGAAGAACTGAGACAGTACG 739
 Db 550 ACTGTAAGGCGAGGTAAGACCACTGCTGAGAGAGGCTTGAAGAACTGAGACAGTACG 609
 QY 740 AACTCAGAAATATTTCCACATGATGATGATGATGATGATGATGATGATGATGATG 799
 Db 610 AACTCAGAAATATTTCCACATGATGATGATGATGATGATGATGATGATGATGATG 669
 QY 800 TGCCCAAGTTTCTCCAGTTAGCAAGTGCACACACTGCGGAGAGTGAAGACATGTGA 859
 Db 670 TGCCCAAGTTTCTCCAGTTAGCAAGTGCACACACTGCGGAGAGTGAAGACATGTGA 729
 QY 860 ATTCTCTGTGAAAACTGATGAGATGAATTTTACCATTTACCATTTCTCATCCCAAG 919
 Db 730 ATTCTCTGTGAAAACTGATGAGATGAATTTTACCATTTACCATTTCTCATCCCAAG 789
 QY 920 CAAAGTGTGGGTCTGATTTACTGCAACACAGAGAGAAAGAACTTTTCTCATACAG 979
 Db 790 CAAAGTGTGGGTCTGATTTACTGCAACACAGAGAGAAAGAACTTTTCTCATACAG 849
 QY 980 GATCAGAGAGGCTCATACACTGAGGCTGATTCAT 1015
 Db 850 GATCAGAGAGGCTCATACACTGAGGCTGATTCAT 885

RESULT 12
 ID ABZ56929 standard; DNA; 1621 BP.
 AC ABZ56929;
 DT 04-APR-2003 (first entry)
 XX
 DE Human AMNH-LP DNA # SEQ ID 1.

XX Human; AMSH-LP; neuroprotective; immunomodulator; cell death;
 KW nervous system; immune system; gene; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 34..1344
 FT /*tag= a
 FT /product= "AMSH-LP"
 PN MO2002102144-A1.
 XX 27-DEC-2002.
 PD 17-JUN-2002; 2002WO-JP006003.
 PF 18-JUN-2001; 2001JP-00184136.
 PR (MISC-) JAPAN SCI & TECHNOLOGY CORP.
 PA Sugamura K, Ishii N;
 PI WPI; 2003-167427/16.
 DR P-PSDB; ABB60075.
 XX Cell death-inducing model non-human animals specifically in nervous or
 PT immune system for study and analysis of cytokine signal transfer
 PT mechanism, applicable in screening promoters or inhibitors.
 PS Claim 20; Page 42-45; 68pp; Japanese.
 CC The invention relates to a cell death-inducible model non-human animal,
 CC characterised by deletion of the AMSH or/and AMSH-LP gene function on the
 CC chromosome. The model animals are applicable in screening promoters or
 CC inhibitors to treat and diagnose diseases in the nervous or immune system
 CC caused by deletion of AMSH and/or AMSH-LP. The current sequence
 CC represents a human AMSH-LP gene fragment
 SQ Sequence 1621 BP; 523 A; 317 C; 336 G; 445 T; 0 U; 0 Other;
 Query Match 22.2%; Score 424.2; DB 7; Length 1621;
 Best Local Similarity 59.5%; Pred. No. 1.3e-105;
 Matches 758; Conservative 0; Mismatches 508; Indels 9; Gaps 2;

QY 431 GAGCTGAAAAGAAAAACAGAGGTAGCACACAGAGAGCAATTTGAAACAGAA 480
 DB 499 TTGATTAAGGCGAAGAAAGCCGATTCTCAATGCGCCAGACGCTAGATCGAG 558
 QY 491 CAGTTCATGCCCTTGAGAGATGATCCGAAACAGAGAGCTAGAAAAAGACGACTGAA 550
 DB 559 CAGTTCTGTTTTTGGAAAGATCAACTCAGAAACAGAGAGTTACCCGAGGTCAAATCGA 618
 QY 551 ATTGTACAGAGATTGGGAAAGTAGACCTGCGCTTAGTGGCCCGCTAGTGGCTGACTTG 610
 DB 619 AGTAGAAACCTCAGAGGCTGTGAGGC--AGATTGATGGAGCCGTTTGTCTGCTTTT 676
 QY 611 GAGAGCCCTCTTAGATGTGTTCCCACTTAACAGTCTCATCCATPACAGCTTCAGAC 670
 DB 677 CCACACACCAAGAACAAATT-----CCTTGTCTAAATGATTTTCAAGATCACTTAATAA 729
 QY 671 TGTACACAACTGTAAAGCCAGCTAAAGCCACTGTGTGAGACAGTCTTGAAACCTGGA 730
 DB 730 AGTAGTCACAACAAATTATGCTAGCCACTCTCTGTAAACAGAGGCTTTAAGCCAGCT 789
 QY 731 GCACTAGCAACTCAGAAAGATATCCCAATGATGATGGCCATGTGGTGGCT 790
 DB 790 GCTACTCTAAGTGTGTTCAAAATTTAGTGTGAAAGACTGAGATGTGATTTTGCA 849
 QY 791 GAGCGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTCCACACTGCCCGGAGATGAG 850
 DB 850 GAAGATCTTGCCACAAATTTCTGCAACTGAGAGATCTAATACGTAGAGAAATAGAA 909
 QY 851 ACATGTGAATTTCTGTGTGAAACACTGATGAGAAATGAATTTCATTACCATGTTCTC 910
 DB 910 ACCTGTGAATCTCTGTGAAACACTGACACATTAATGAATTAATTAATTAATTAAT 969
 QY 911 ATCCCAAGCAAGTCTGGGTCTGATTACTGCAACAGAGAGCAAGAAAGAACTTTTC 970
 DB 970 GTCCCAAGACGTCTGGGACCAAGCTATTTGATGATGAGAAATGAGAAATTAATTC 1029
 QY 971 CTGATACAGATCAGAGGAGGCTCATCACATGGGCTGATTCATCTACCCACACAG 1030
 DB 1030 AATGTACAGATCAATGATCTCTCACTTAGATGATCCATCACATCCCACTCAA 1089
 QY 1031 ACCGCTTCTCTCCAGTGTGACCTAAGACCTACCTGCTTACAGATGATTTGCCA 1090
 DB 1090 ACTGCAATTTTATCCAGCTGTGATTTCTCACTCACTGTCTTATCACTCAATGTTGCCA 1149
 QY 1091 GAGTCAGTACCATGTTTGTCTCCCAAGTTCCAGAAATCGAATTTCTTAAACTACT 1150
 DB 1150 GAGGCCATTTGCCATTTGTTGTCAACCAAGACATTAAGACACTGGAATCTTCAAGGCTCAC 1209
 QY 1151 GACCATGACTAGAGAGATTTCTCTGTGCGCCAGAAAGATTTGATCCACACAGCAAG 1210
 DB 1210 AATGCTGGCATCTTGAGGTTTCTGTGTAAAAAAAGGCTTTCATCCACACACCAAG 1269
 QY 1211 GATCCACTCTGTTCTGTAGTGCAGCCAGCTGATGTGTGAGAGAGCACTGACATC 1270
 DB 1270 GAGCCAGGCTTTGATGATATGCAAAACATGTGTTGTTAAAGACATTAATAATTTG 1329
 QY 1271 ACAGACTTTCGATGA 1285
 DB 1330 TTGATCTAGGTGA 1344
 RESULT 13
 ABZ56931
 ID ABZ56931 standard; DNA; 1392 BP.
 XX ABZ56931;
 AC 04-APR-2003 (first entry)
 XX Mouse AMSH-LP DNA # SEQ ID 5.
 DE Mouse; AMSH-LP; neuroprotective; immunomodulator; cell death;
 KW nervous system; immune system; gene; ds.


```

XX OS Mus musculus.
XX FH Key Location/Qualifiers
XX FT CDS 72..1382
XX FT /*tag= a
XX FT /product= "AMSH-LP"
XX PN WO2002102144-A1.
XX PD 27-DEC-2002.
XX PF 17-JUN-2002; 2002WO-JP006003.
XX PR 18-JUN-2001; 2001JP-00184136.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Sugamura K, Ishii N;
XX PI MPI: 2003-167427/16.
XX DR P-PSDB; ABP6007.
XX PT Cell death-inducing model non-human animals specifically in nervous or
XX PT immune system for study and analysis of cytokine signal transfer
XX PT mechanism, applicable in screening promoters or inhibitors.
XX PS Claim 26; Page 54-57; 68pp; Japanese.
XX CC The invention relates to a cell death-inducible model non-human animal,
XX CC characterised by deletion of the AMSH or/and AMSH-LP gene function on the
XX CC chromosome. The model animals are applicable in screening promoters or
XX CC inhibitors to treat and diagnose diseases in the nervous or immune system
XX CC caused by deletion of AMSH and/or AMSH-LP. The current sequence
XX CC represents a mouse AMSH-LP gene fragment
XX SQ Sequence 1392 BP; 408 A; 322 C; 332 G; 330 T; 0 U; 0 Other;
Query Match 21.9%; Score 418.2; DB 7; Length 1392;
Best Local Similarity 58.7%; Pred. No. 5.4e-104;
Matches 748; Conservative 0; Mismatches 518; Indels 9; Gaps 1;
QY 11 ATGCTGACCCATGAGATGTGAGCTCCGCCGAGACGGGGTGAAGGCTCTCTCCAG 70
DB 117 ATGCTGACCATACAGATGTTCTCTAAGTCAAGAGAGCGGGTCCGCCCTTAAGCAA 176
QY 71 CTGGGTAGTGGCGTAGAGGTGATGAAGACATCCACCCGTCGGTACTTCCGCTCGA 130
DB 177 CTGGGTCTATATCTCCATTAATGAAGATATCACCCGCGCTACTTCAAGGTCGGA 236
QY 131 GTTGAATTTATCCGATGGCATTCATTACTCTGAGAGAGCAATTTGAACATGCTTC 190
DB 237 GTGAAATGAAAGGATGCACTCTGTATTGGAAGAGGAAACCTGAAAAATGCTTT 296
QY 191 ATCCCTATTAACAAGTATATCAAGCTCTTTATTGAGAACTACCAAAACATCGAGATTAC 250
DB 297 GTTCTTTTAACTAATTTAATACGTTATTGTAGAAATCTTCCAGCACCAGATTAC 356
QY 251 AAATCTGCTGCTATTCCTGAAAGAAAGACACAGTAAAGAAATTTAAAGAGATTGCAATT 310
DB 357 CAGCAGTGTGAGTTCAGAGAGAGAGATATTATGAAGAACTGAAAGAGATTCGCTTC 416
QY 311 CCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370
DB 417 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
QY 371 TATATGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
DB 477 TATTTGCAAAAGTAAACAAATATATAAGCCGAAATTTCTCAAAAAGTTGGAACATCAGAGA 536
QY 431 GAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 490
DB 537 CTGATAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596

```

```

QY 491 CAGTTCATGCTCTCGAGAGATGATCCGAAACCAAGAGCTAGAAAAAGAGCACTGAAA 550
DB 597 CAGTTCCTGTTTGTGAAGCAACTCAAGAGAGAGAGTGGCTCGAGGCCAGATTGCA 656
QY 551 ATTGTACAGAGATTGGGAAGAGTGAACCTTCGCTTACGTGGCCCGCTAGTGCCCTGACTTG 610
DB 657 GGTGAGAGACTCTCCGGGTGTGTCTGAGAGAGCTGACGAAAGTGGCTGTCTGCTTTCC 716
QY 611 GAGAAAGCCCTCTTAGATGTGTCCCACTTTAACAGTCTTCATCCATACAGCTTTCAGAC 670
DB 717 ACCCAGCAGAGCACTCTCTAGAGAGAGAGATTTGCAAT-----CACCTCATATAA 767
QY 671 TGTACACACTGTGAAGCCAGCTTAAGCCACTGTGTGTGACAGTCTCTGAACCTGGA 730
DB 768 AGTGTAGAGCAATTTGCTTAATCTCTCTCCGTAAGAGAGAGAGAGAGAGAGAGAGAG 827
QY 731 GCACAGACCACTAGAAAGTATCCCACTATGATGATTTGCCCATGTGTGTGCT 790
DB 828 GCCACCTGAGTGTCTTCAAGATTTGGTGTGTAAGGCGCTGAGGTGTGTATTTATCA 887
QY 791 GGGCGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTCCAAACTCTCCCGGAGGTGAG 850
DB 888 AGAGATCTTTGGCATTAATTTCTGTGCTGGCTGACCTTAACAGAGAGAGAGAGAGAG 947
QY 851 ACATGTGGAATTTCTGTGAGAAACTGATGAGATGATTTACCATTAACCATTTCTC 910
DB 948 ACCTGTGAGATCTCTGTGAGAAACTGACACAAATGAATTCACCATTAATCTGTGTT 1007
QY 911 ATCCCAAGCAAGTGTGGGTGTGATTAATCTGCAACACAGAGAAAGAAAGAACTTTTC 970
DB 1008 GTGCCAAGAGAGTGTGTGGCCAGAGATTAATGCAATGTGAGAGAGAGAGAGAGAG 1067
QY 971 CTCATACAGAGATACAGAGGCTTATCACTAGGCTGTGATTTACTACACCCACAG 1030
DB 1068 AATGTTGAGATCAACATGTGTCTCTCACTGAGATGATCAACACCCACAG 1127
QY 1031 ACCGCTTCTCTCAGTGTGACCTTACACACTACTCTCTTACAGATGTTGCA 1090
DB 1128 ACGCATTCCTTCCAGTGTGATCTCCACACTACTCTCTTACAGATGTTGCA 1187
QY 1091 GAGTACAGTACCATTTGTTGCTCCCAAGTTCCAGAGAACTGATTTTAACTACT 1150
DB 1188 GAGGCATTTGCTGTGTGTGTCTCCCAAGATTAAGACACCGGATTTGAGGCTCAC 1247
QY 1151 GACCATGAGCTAGAGAGATTTCTCTGTGCGCAAGAGATTTCAATCCACAGCAAG 1210
DB 1248 AACGCTGATCTGTGAGGTTTCTACTGTAAAGAGGCTTCCATCTTACAGCAAG 1307
QY 1211 GATCCACTCTGTCTGTAGCTGACGACGATGTTGTGACAGAGAGAGAGAGAG 1270
DB 1308 GACCCCAAGCTGTTCAGTATGACAGCCATGTGTAGTAAGAGATTAAGAAACATGTG 1367
QY 1271 ACAGACCTTCAGTGA 1285
DB 1368 TTGATCTGAGGTGA 1382

```

RESULT 14
ABZ56930 standard; DNA; 1428 BP.

ABZ56930;
04-APR-2003 (first entry)
Mouse AMSH-LP DNA # SEQ ID 3.
Mouse; AMSH-LP; neuroprotective; immunomodulator; cell death;
nervous system; immune system; gene; ds.
Mus musculus.

FH Key Location/Qualifiers
 FT CDS 108..1418
 FT /*tag= a
 FT /product= "AMSH-LP"
 XX
 XX MO2002102144-A1.
 XX
 XX 27-DEC-2002.
 XX
 XX 17-JUN-2002; 2002MO-JP006003.
 XX
 XX 18-JUN-2001; 2001JP-00184136.
 XX
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 XX Sugamura K, Ishii N;
 XX
 XX MPI: 2003-167427/16.
 XX P-PSDB; ABB60076.
 XX
 XX Cell death-inducing model non-human animals specifically in nervous or
 PT immune system for study and analysis of cytokine signal transfer
 PT mechanism, applicable in screening promoters or inhibitors.
 XX
 XX Claim 23; Page 48-51; 68pp; Japanese.
 XX
 XX The invention relates to a cell death-inducible model non-human animal,
 CC characterised by deletion of the AMSH or/and AMSH-LP gene function on the
 CC chromosome. The model animals are applicable in screening promoters or
 CC inhibitors to treat and diagnose diseases in the nervous or immune system
 CC caused by deletion of AMSH and/or AMSH-LP. The current sequence
 CC represents a mouse AMSH-LP gene fragment
 XX
 XX Sequence 1428 BP; 412 A; 336 C; 345 G; 335 T; 0 U; 0 Other;
 SO
 Query Match 21.9%; Score 418.2; DB 7; Length 1428;
 Best Local Similarity 58.7%; Pred. No. 5.5e-104;
 Matches 748; Conservative 0; Mismatches 518; Indels 9; Gaps 1;

Db 633 CAGTTCCTGTTTTTTGAAGCAACTCAAGAGAGAGATTGGCTGAGGCCAGATTGCA 692
 Qy 551 ATTGTACAGAGATTGGGAAAGTAGACCTGGGCTAGAGGCCGCTAGTGCCTGACTTG 610
 Db 693 GGTGAGGACTCTCCGGTGTGTCTGAGCAGACTGACGGAAGTGCGCTGTCTGTTC 752
 Qy 611 GAGAGCCCTCCTTAGTGTGTGTCCCACTTAACTCTACATCATACGCTTACGAC 670
 Db 753 ACCCAGCAGAGCACTCTGAGGAACGCAATTGCAAGT-----CACCTCATMAA 803
 Qy 671 TGTACACAACCTGTAAGGCCAGCTAAGCCACCTGTGGTGACAGTCTTGAACCTGGA 720
 Db 804 AGTGATGGAAGCAATTTCCGTAACACTCTCCCTGTAACAGGGCCCTTAAGCCAGCA 853
 Qy 731 GCACGTAGCAACTGAGAAATTTCCCAATGATGATGTGGCCATGTGTGTGCTT 790
 Db 864 GCCACCCGTAGTGTCTGCAAAATTTGTGTGAAGGGCTGAGTGTGATTATCA 923
 Qy 791 GGCAGGCTGTGCCCAAGTTTCTCCAGTTAGCCAGTGCACACTGCCGGGAGTGAG 850
 Db 924 AGAGATCTTGGCATTAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
 Qy 851 ACATGTGAATTTCTGTGTGAAAATGATGAGGAATGAATTACATTAACCATGTCTC 910
 Db 984 ACCGTGGGATCTCTGTGAAAATGACACACATGAATTCACTTACTCATGTGCTT 1043
 Qy 911 ATCCCAAGAAAGTCTGGTGTCTGATTTACTGCAACAGAGAAAGCAAGCACTTTC 970
 Db 1044 GTCCAAAGAGTCTGTGGCCCAAGCTATGTGATGTGAAGAACTCGAAGATTATTC 1103
 Qy 971 CTATACAGATCAGAGAGGCGCTCATCACTGGGCTGATTCATCACTACCCACAG 1030
 Db 1104 AATGTCAGGATCAACATGCTCTCCACACTGGATGATCATACATCAACCCAGCAA 1163
 Qy 1031 ACCGCTTTCTCTCCAGTGTGACCTACACTCACTGCTCTTACCAATGATGTGCA 1090
 Db 1164 ACCGCACTCTCTGCAAGTGTGATCTCCACTCACTGCTCTTACCAAGCTTATGTGCA 1223
 Qy 1091 GAGTCAGTAGCCATTTTGTGCTCCCAAGTTCAGGAAGCTGGAATCTTAACTA 1150
 Db 1224 GAGGCCATTCATTTGTGTGTCTCCCAAGCAATMAAGACCCGCACTTTCAGCTCAC 1283
 Qy 1151 GACCATGACTAGAGAGATTTCTTCTGTGCGCAAGAAAGATTATCATCACAAGCAAG 1210
 Db 1284 AACGCTGCAATCTTGAAGTTCTTCTTAAGAAAGAGGCTTCATCTCACAACMAAG 1343
 Qy 1211 GATCCACTCTGTGTGTGAGCTGCAAGCCAGCACTGTGTGTGAACAGACAGTGCATC 1270
 Db 1344 GACCCCAAGCTTTCAATATATGACGCAATGTGTATGAAGACATATAAAACAATGTG 1403
 Qy 1271 ACAGACCTTGATGA 1285
 Db 1404 TTGATCTGAGGTGA 1418
 RESULT 15
 AA159567
 ID AA159567 standard; cDNA; 4052 BP.
 XX
 XX AA159567;
 XX
 XX 22-OCT-2001 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 1770.
 XX
 XX Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 XX
 XX Homo sapiens.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: August 23, 2004, 10:00:58 ; Search time 7421 Seconds

(without alignments)
1155.531 Million cell updates/sec

Title: US-09-831-452-2

Perfect score: 1910

Sequence: 1 cttgtctctgctgtctgtacc.....ataaagtaacaattactta 1910

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : GenBank1.*
1: gb_ba.*
2: gb_bg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_in.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pin.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1910	100.0	1910	6 E35541	E35541 Protein AWS
2	1910	100.0	1930	6 HSU073522	U73522 Homo sapien
3	1910	100.0	2003	6 AX100575	AX100575 Sequence
4	1908.4	99.9	1923	9 BC007662	BC007662 Homo sapi
5	1366	71.0	1462	9 AF052155	AF052155 Homo sapi
6	1043	54.6	1043	6 BD264634	BD264634 Compositi
7	1043	54.6	1043	6 AR238037	AR238037 Sequence
8	1043	54.6	1043	6 AR238037	AR238037 Sequence
9	1043	54.6	1043	6 AR238037	AR238037 Sequence
10	1043	54.6	1043	6 AR238037	AR238037 Sequence
11	1040	54.5	1041	6 BD264633	BD264633 Compositi
12	1040	54.5	1041	6 AR238036	AR238036 Sequence
13	1040	54.5	1041	6 AR238036	AR238036 Sequence
14	1040	54.5	1041	6 AR238036	AR238036 Sequence
15	1040	54.5	1041	6 AR238036	AR238036 Sequence
16	889.8	46.6	1544	10 AY083159	AY083159 Rattus no
17	889.8	46.6	1735	10 BC061711	BC061711 Rattus no
18	879.8	46.1	1384	6 E35542	E35542 Protein AWS
19	879.8	46.1	2094	10 BC025111	BC025111 Mus muscu
20	879.8	46.1	2104	10 AB010123	AB010123 Mus muscu
21	879.8	46.1	2122	10 BC003497	BC003497 Mus muscu
22	879.8	46.1	2172	10 BC006939	BC006939 Mus muscu
23	683	35.8	145140	2 AC073577	AC073577 Homo sapi
24	683	35.8	173658	9 AC073046	AC073046 Homo sapi
25	624.4	32.7	56315	2 AC137140	AC137140 Homo sapi
26	519.2	27.2	2457	5 AY398309	AY398309 Dario rex
27	514	26.9	2455	5 BC055512	BC055512 Dario rex
28	442.8	23.2	1828	5 BC061390	BC061390 Silurana
29	425.8	22.3	2010	6 BC010846	BC010846 Homo sapi
30	424.2	22.2	1601	9 AB010120	AB010120 Homo sapi
31	424.2	22.2	1621	6 BD186116	BD186116 Apoptosis
32	424.2	22.2	1621	6 BD187452	BD187452 Apoptosis
33	424.2	22.2	2006	9 AK056086	AK056086 Homo sapi
34	422.6	22.1	2206	9 AK092946	AK092946 Homo sapi
35	418.2	21.9	1392	6 BD186118	BD186118 Apoptosis
36	418.2	21.9	1392	6 BD187454	BD187454 Apoptosis
37	418.2	21.9	1428	6 BD186117	BD186117 Apoptosis
38	418.2	21.9	1428	6 BD187453	BD187453 Apoptosis
39	418.2	21.9	1690	10 AB066213	AB066213 Mus muscu
40	418.2	21.9	2008	10 AB066211	AB066211 Mus muscu
41	415	21.7	1717	10 BC018343	BC018343 Mus muscu
42	413	21.6	4052	9 AB037794	AB037794 Homo sapi
43	407	21.3	1973	6 AX082297	AX082297 Sequence
44	369.4	19.3	143325	2 AL137180	AL137180 Homo sapi
45	369.4	19.3	148785	2 AC105412	AC105412 Homo sapi

ALIGNMENTS

RESULT 1
E35541 LOCUS E35541 1910 bp DNA linear PAT 31-JAN-2002
DEFINITION Protein AWS and CDNA thereof.
ACCESSION E35541
VERSION E35541.1 GI:18624562
KEYWORDS JP 2000139469-A/1
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1910)
AUTHORS Sugamura, K. and Tanaka, N.
TITLE Protein AWS and CDNA thereof
JOURNAL Patent: JP 2000139469-A 1 23-MAY-2000;

COMMENT SCIENCE & TECH AGENCY
OS Homo sapiens (human)
PN JP 2000139469-A/1
PD 23-WAY-2000
PR 12-NOV-1998 JP 1998322674
PI KAZUO SUGAMURA, NOBUYUKI TANAKA
PC C12N15/09, C07K14/52, C07K16/24//C12N1/21, C12P21/02, (C12N15/09,
PC C12N1/21),
PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), C12N15/00, (C12N15/00,
PC C12R1:91)
CC
FH Key Location/Qualifiers
FT CDS 11.1282.
FEATURES
source 1.1910
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1910; DB 6; Length 1910;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTGCTCCTGATCTGACATGAGATGAGCTCCGCGGAGAGACCGGGTGAAGGC 60
DB 1 CTTGCTCCTGATCTGACATGAGATGAGCTCCGCGGAGAGACCGGGTGAAGGC 60
QY 61 TCTCTCCAGCTGCGGTAGTGGGTAGAGTGAATGAAGACATTCACCCCGGTACTT 120
DB 61 TCTCTCCAGCTGCGGTAGTGGGTAGAGTGAATGAAGACATTCACCCCGGTACTT 120
QY 121 CCGCTCTGAGATGATATCCGAATGCGATCATTACTCTGAGAGAGCAACATTTGA 180
DB 121 CCGCTCTGAGATGATATCCGAATGCGATCATTACTCTGAGAGAGCAACATTTGA 180
QY 121 CCGCTCTGAGATGATATCCGAATGCGATCATTACTCTGAGAGAGCAACATTTGA 180
DB 121 CCGCTCTGAGATGATATCCGAATGCGATCATTACTCTGAGAGAGCAACATTTGA 180
QY 181 ACATGCTTCATCTCTATTAACAGTATATCACTCTTATTTAGAGAACTACCAAAACA 240
DB 181 ACATGCTTCATCTCTATTAACAGTATATCACTCTTATTTAGAGAACTACCAAAACA 240
QY 181 ACATGCTTCATCTCTATTAACAGTATATCACTCTTATTTAGAGAACTACCAAAACA 240
DB 181 ACATGCTTCATCTCTATTAACAGTATATCACTCTTATTTAGAGAACTACCAAAACA 240
QY 241 TCGAGATTACAAATCTGCTGCTATCTCTGAAAGAAAGACACAGTAAAGAAATTAAAGA 300
DB 241 TCGAGATTACAAATCTGCTGCTATCTCTGAAAGAAAGACACAGTAAAGAAATTAAAGA 300
QY 301 GATTGCAATTCGCAAGAGAGAGCTGAGAGGAGAGAGCTGTAAACGATATACCAAGA 360
DB 301 GATTGCAATTCGCAAGAGAGAGCTGAGAGGAGAGAGCTGTAAACGATATACCAAGA 360
QY 361 ATATACAGATATATGAG 420
DB 361 ATATACAGATATATGAG 420
QY 421 CATTCAG 480
DB 421 CATTCAG 480
QY 481 GGAACAG 540
DB 481 GGAACAG 540
QY 541 GCGACTGAAATTTGTAAG 600
DB 541 GCGACTGAAATTTGTAAG 600
QY 601 GCGTGAAG 660
DB 601 GCGTGAAG 660
QY 661 GCGTGAAG 720
DB 661 GCGTGAAG 720

QY 721 GAAACTGAG 780
DB 721 GAAACTGAG 780
QY 781 GGTGCTGCTGAG 840
DB 781 GGTGCTGCTGAG 840
QY 841 GGAAGTGAAG 900
DB 841 GGAAGTGAAG 900
QY 901 CCATGTTCTCATCCCAAG 960
DB 901 CCATGTTCTCATCCCAAG 960
QY 961 AGAATTTTCTCATACAG 1020
DB 961 AGAATTTTCTCATACAG 1020
QY 1021 CCCGACAG 1080
DB 1021 CCCGACAG 1080
QY 1081 GATGTTGCCAG 1140
DB 1081 GATGTTGCCAG 1140
QY 1141 TAAACTTAACAG 1200
DB 1141 TAAACTTAACAG 1200
QY 1201 ACACAG 1260
DB 1201 ACACAG 1260
QY 1261 AGTGAACATACAG 1320
DB 1261 AGTGAACATACAG 1320
QY 1321 CCATGTTCTCATCCCAAG 1380
DB 1321 CCATGTTCTCATCCCAAG 1380
QY 1381 TAGATGATGAAAG 1440
DB 1381 TAGATGATGAAAG 1440
QY 1441 AAGAAATTAAG 1500
DB 1441 AAGAAATTAAG 1500
QY 1501 AACTGTAACAG 1560
DB 1501 AACTGTAACAG 1560
QY 1561 GGTATTAATGAAG 1620
DB 1561 GGTATTAATGAAG 1620
QY 1621 CTCTGAGTATCTCTATTAATTTCTCTATTAATTTCTCTATTAATTTCTCTATTAATTT 1680
DB 1621 CTCTGAGTATCTCTATTAATTTCTCTATTAATTTCTCTATTAATTTCTCTATTAATTT 1680
QY 1681 CAATTAAGGAG 1740
DB 1681 CAATTAAGGAG 1740
QY 1741 TGGTTATGAG 1800
DB 1741 TGGTTATGAG 1800
QY 1801 ACCAGATGTTGGCTTAACAG 1860
DB 1801 ACCAGATGTTGGCTTAACAG 1860

```

Db      1801 ACCGATGTTGGTACACATCCCGAAGATGATTTTGCAGGAATTATTGTATTTA 1860
QY      1861 ATAAATATTCAGATATTTTCCCTTACATTAAGTAAAGTAACTTA 1910
Db      1861 ATAAATATTCAGATATTTTCTCTACAAATTAAGTAACTTA 1910

RESULT 2
HSU73522      1930 bp      mRNA      linear      PRI 29-JUN-1999
LOCUS      HSU73522
DEFINITION      Homo sapiens AMSH mRNA, complete cds.
ACCESSION      U73522
VERSION      U73522.1
KEYWORDS      GI:4098123
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1930)
AUTHORS      Tanaka,N., Kaneo,K., Asao,H., Kasai,H., Endo,Y., Fujita,T.,
Takeshita,T. and Sugamura,K.
ABSTRACT      Possible involvement of a novel STAM-associated molecule 'AMSH' in
intracellular signal transduction mediated by cytokines
J. Biol. Chem. 274 (27), 19129-19135 (1999)
JOURNAL      MEDLINE
PUBMED      10383417
REFERENCE      2 (bases 1 to 1930)
AUTHORS      Tanaka,N., Kaneo,K., Kasai,H., Takeshita,T. and Sugamura,K.
ABSTRACT      Direct Submission
JOURNAL      Submitted (07-OCT-1996) Microbiology, Tohoku University School of
Medicine, 2-1 Seiryō-machi Aobaku, Sendai 980-77, Japan
FEATURES
Source
1..1930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="PHA-PBL"
11..1285
/codon_start=1
/product="AMSH"
/protein_id="A4D05037.1"
/db_xref="GI:4098124"
/translation="MSDHDVSLPPEEDVRALSOLGSAVENVEDIPPRVFRSGVFI
RMAISVEENINIEAFILVYKYLITLPEKPKRDKYSNIIPKKTIVKUKLAIAPK
ABEIKALIKRYTKEITETVBEKKKEBEELARMAIOOELEKQVADQKQOOLQK
QFHAPEMIRNQELEKRLKI VQEPKGVDELGLPLVDELKESLDVFPITLVSSIQP
SDGHTVPAPKPVVDRSLKPGALSNSIS.PITDGLRHVVVPERLCPVLOLASANTA
RGVETGILCGKMLNEFTITVHLIPKOSAGSDYCNTEENEEFLIQQGLITLGM
HTHPTQAFPLSSVDLTHGCSYOMLPEKSAIVCSPKFQETGFKLIDHGLEISCRQ
KGFHPHSDPPLFSCSHVTVVRAVYITDLR"
ORIGIN
Query Match      100.0%; Score 19.0; DB 9; Length 1930;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1      CTGGGTCTATGTCACCATGAGATGTGAGCTCCGCCGGAAGACGGGTGAGGCG 60
Db      1      CTGGGTCTATGTCCTACCATGAGATGTGAGCTCCGCCGGAAGACGGGTGAGGCG 60
QY      61      TCTCTCCCACTGGGTAGTGGGTAGAGTGAATGAAGCATTCACCCCGTCGACTT 120
Db      61      TCTCTCCCACTGGGTAGTGGGTAGAGTGAATGAAGCATTCACCCCGTCGACTT 120
QY      121      CCGGTCTGAGGTGAGATTATCCGAATGCGATCATTACTCTGAGGAAAGCAATGGA 180
Db      121      CCGGTCTGAGGTGAGATTATCCGAATGCGATCATTACTCTGAGGAAAGCAATGGA 180
QY      181      ACATGCTTATCTCTATTAACAAGTATATCAGCTCTTTATTGAGAAATACCAAAAC 240
Db      181      ACATGCTTATCTCTATTAACAAGTATATCAGCTCTTTATTGAGAAATACCAAAAC 240

```

```

QY      241      TCGAGATTACAAATCTGTCATTCCTGAAAAAGAAAGACAGATGAATAATTAAAGA 300
Db      241      TCGAGATTACAAATCTGTCATTCCTGAAAAAGAAAGACAGATGAATAATTAAAGA 300
QY      301      GATTGATTTCCCAAGACAGAAAGCTGAAGCAGAGCTGTTAAACGATATACCAAGA 360
Db      301      GATTGATTTCCCAAGACAGAAAGCTGAAGCAGAGCTGTTAAACGATATACCAAGA 360
QY      361      ATATACAGATATTAATGAAGAAAGAAAGAGAGAGAGGAAATGGCCCGGAACATGCG 420
Db      361      ATATACAGATATTAATGAAGAAAGAAAGAGAGAGAGGAAATGGCCCGGAACATGCG 420
QY      421      CATCCAGAAAGCTGAAAAAGAAAAACAGAGGTAGACAAACGAAACAGACAGCAATT 480
Db      421      CATCCAGAAAGCTGAAAAAGAAAAACAGAGGTAGACAAACGAAACAGACAGCAATT 480
QY      481      GGAACAGAAACAGTTCCATGCTTGGAGAGATGATCCGGAACGAGACTTGAAGAAAA 540
Db      481      GGAACAGAAACAGTTCCATGCTTGGAGAGATGATCCGGAACGAGACTTGAAGAAAA 540
QY      541      GCGACTGAAAAATTTGACAGAGATTGGAGAGGTAGACCTGAGCTAGTGCCCGCTAGT 600
Db      541      GCGACTGAAAAATTTGACAGAGATTGGAGAGGTAGACCTGAGCTAGTGCCCGCTAGT 600
QY      601      GCTGACTTGGAGAAAGCTTCTTGAATGTGTTCCCACTTAACATGTCTCATATACA 660
Db      601      GCTGACTTGGAGAAAGCTTCTTGAATGTGTTCCCACTTAACATGTCTCATATACA 660
QY      661      GCCTTAGACTGTACACAACTGTAGAGCCACTTAAGCAGCTGGTGGAGACAGGCTCT 720
Db      661      GCCTTAGACTGTACACAACTGTAGAGCCACTTAAGCAGCTGGTGGAGACAGGCTCT 720
QY      721      GAAACCTGAGACCTGAGCACTAGCAAGAAATATCCCAATCGATGATTTGGCCATGT 780
Db      721      GAAACCTGAGACCTGAGCACTAGCAAGAAATATCCCAATCGATGATTTGGCCATGT 780
QY      781      GGTGGTGGCTGGGGGGGCTGTGGCCACAGTTTCCAGTTAGGCAATGCGCAACATGCGCG 840
Db      781      GGTGGTGGCTGGGGGGGCTGTGGCCACAGTTTCCAGTTAGGCAATGCGCAACATGCGCG 840
QY      841      GGGAGTGGACATGTGGAATCTCTGTGAAAACTGATGAGAAATGAATTAACATTAC 900
Db      841      GGGAGTGGACATGTGGAATCTCTGTGAAAACTGATGAGAAATGAATTAACATTAC 900
QY      901      CCAATGTTCTATCCCAAGCAAGTGTGCTGTGATTAATCTGCAACACAGAGACGAAGA 960
Db      901      CCAATGTTCTATCCCAAGCAAGTGTGCTGTGATTAATCTGCAACACAGAGACGAAGA 960
QY      961      AGAACTTTTCCCATACAGAGATCAGAGAGGCTCATCACTGGGCTGGATTCATATCA 1020
Db      961      AGAACTTTTCCCATACAGAGATCAGAGAGGCTCATCACTGGGCTGGATTCATATCA 1020
QY      1021      CCCCAACAGACCGGGTCTCTCCAGTGTGCACTACACTACTCTCTTACCAAGAT 1080
Db      1021      CCCCAACAGACCGGGTCTCTCCAGTGTGCACTACACTACTCTCTTACCAAGAT 1080
QY      1081      GATGTTGCCAGAGTACAGATCCATGTTTGTCTCCCAAGTTCAGGAACCTGATTTCT 1140
Db      1081      GATGTTGCCAGAGTACAGATCCATGTTTGTCTCCCAAGTTCAGGAACCTGATTTCT 1140
QY      1141      TAAACTTAATGACATGACATTAAGAGAAATTTCTCTGTGCGCCAGAAAGATTTATCC 1200
Db      1141      TAAACTTAATGACATGACATTAAGAGAAATTTCTCTGTGCGCCAGAAAGATTTATCC 1200
QY      1201      AACAAGCAAGATCCACCTGTGTTGTAGCTGACAGCAGGATCTGTGGAGACAGAGC 1260
Db      1201      AACAAGCAAGATCCACCTGTGTTGTAGCTGACAGCAGGATCTGTGGAGACAGAGC 1260
QY      1261      AGTGACATCAACAGACTTTCATGAGAGCTTTGAGTCCAAACCTTCCAAAGAACAA 1320
Db      1261      AGTGACATCAACAGACTTTCATGAGAGCTTTGAGTCCAAACCTTCCAAAGAACAA 1320
QY      1321      CCATATCAGTACTGTAGCCCTTAATTAAAGCTTTCTAGAAAAGCTTTGAGAGTTTTC 1380

```



```

Db      1013 AGAATCTTCTCATACAGAGATCAGCAGGCGCTCATCAGCTGGGCTGATTCATCTCA 1072
Qy      1021 CCCGACAGACCGGGTCTCTCCAGTGTGACCTTACACTGCTCTTACCGAT 1080
Db      1073 CCCACAGACGCGGGTCTCTCCAGTGTGACCTTACACTGCTCTTACCGAT 1132
Qy      1081 GATGTTCAGAGTCAAGTCAATGTTGCTCCCGCAAGTTCAGAGAACTGATCTT 1140
Db      1133 GATGTTCAGAGTCAAGTCAATGTTGCTCCCGCAAGTTCAGAGAACTGATCTT 1192
Qy      1141 TAACTACTGACCTGACCTAGAGAGATTTCTCTGTGCCAGAAAGATTTATCC 1200
Db      1193 TAACTACTGACCTGACCTAGAGAGATTTCTCTGTGCCAGAAAGATTTATCC 1252
Qy      1201 ACACAGCAGAGATCCACTCTGTCTGTAGCTGCAGCAGGACGTCGTGTGTGACAGAGC 1250
Db      1253 ACACAGCAGAGATCCACTCTGTCTGTAGCTGCAGCAGGACGTCGTGTGTGACAGAGC 1312
Qy      1261 AGTGAACATCAACAGCCTTGATGAGCGTTTGATCCAAACCTTCCAAAGAACAAAA 1320
Db      1313 AGTGAACATCAACAGCCTTGATGAGCGTTTGATCCAAACCTTCCAAAGAACAAAA 1372
Qy      1321 CCATTCAGTGTAGTCTGAGCCCTTAAATTAAGCTTTCTAAGAAAGCTTTGGAAGTTTGG 1380
Db      1373 CCATTCAGTGTAGTCTGAGCCCTTAAATTAAGCTTTCTAAGAAAGCTTTGGAAGTTTGG 1432
Qy      1381 TAGATAGTAAAGGGGGGACATCACTGAGAAAGAGTGAATTTGTATTCAGTTTGA 1440
Db      1433 TAGATAGTAAAGGGGGGACATCACTGAGAAAGAGTGAATTTGTATTCAGTTTGA 1492
Qy      1441 AAGAAATACCTGACATTTTTTTAGGCAAGTCAAGAAAGAGATGTGTACCAAAAGC 1500
Db      1493 AAGAAATACCTGACATTTTTTTAGGCAAGTCAAGAAAGAGATGTGTACCAAAAGC 1552
Qy      1501 AACTGTAACTCAGAAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGAAAGAT 1560
Db      1553 AACTGTAACTCAGAAATTAAGTACTCAGAAATTAAGTACTCAGAAATTAAGAAAGAT 1612
Qy      1561 GGTATATGAACCCCATATACCTTCTGTGATTCACCAATTTGTAACATTTTTTC 1620
Db      1613 GGTATATGAACCCCATATACCTTCTGTGATTCACCAATTTGTAACATTTTTTC 1672
Qy      1621 CTCACAGTATCTTCTTAATTTCTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 1680
Db      1673 CTCACAGTATCTTCTTAATTTCTCTTAATTTCTTAATTTCTTAATTTCTTAATTTCT 1732
Qy      1681 CAATTAAGGGCATCTGTGAGAAATTTGAGAGCCATTTAGAAATCTTTGGATTTCTG 1740
Db      1733 CAATTAAGGGCATCTGTGAGAAATTTGAGAGCCATTTAGAAATCTTTGGATTTCTG 1792
Qy      1741 TGTGTTATGGCAATATGATGAGACTTATTACTGGGGTGAAGGACAGCTTACTCATTG 1800
Db      1793 TGTGTTATGGCAATATGATGAGACTTATTACTGGGGTGAAGGACAGCTTACTCATTG 1852
Qy      1801 ACCAGATTTGTTGGCTAACACATCCGAGAGATTTTGTGAGAAATTAATTTATTTA 1860
Db      1853 ACCAGATTTGTTGGCTAACACATCCGAGAGATTTTGTGAGAAATTAATTTATTTA 1912
Qy      1861 ATAAATATTTTCAAGATATTTTCTCTTACCAATTAAGTAACTTAATTA 1910
Db      1913 ATAAATATTTTCAAGATATTTTCTCTTACCAATTAAGTAACTTAATTA 1962

```

```

RESULT 4
LOCUS   BC007682      1933 bp      mRNA      linear      PRI 03-OCT-2003
DEFINITION Homo sapiens associated molecule with the SH3 domain of STM, mRNA
ACCESSION BC007682
VERSION   BC007682.1 GI:14043381
KEYWORDS  MGC.
SOURCE    Homo sapiens (human)

```

ORGANISM

REFERENCE
AUTHORS

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1933)
Strausberg, R.L., Feilgold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavert, T.L.,
Scheetz, T.E., Brownstein, M.U., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Muliani, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schmeckel, A., Schein, J.E., Jones, S.J., Skalska, U., Small, D.E.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1933)
Strausberg, R.
Direct Submission
Submitted (11-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Louis W. Staudt, M.D., Ph.D.
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@ncl.nih.gov
Ahter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granter, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Scantirip, S., Thomas, P.D., Touchman, J.W.,
Tsurgoev, C., Vogt, J.D., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNU at: <http://image.llnl.gov>
Series: IRM plate: 12 Row: 1 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 17738303.
Location/Qualifiers
1. 1933

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="WGC:3546 IMAGE:3619837"
/tissue_type="lymph, Burkitt lymphoma"
/clone_lib="NIH MGC 8"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
1. 1933
/gene="AMSH"
/db_xref="locusID:10617"
/db_xref="MIM:606247"

```



```

Db      1747 TGGTTATGCAATATGANTGAGCTTATTAAGTGGGGTGGAGGACAGCTTACCTCATTG 1806
QY      1801 ACCAGATTGTTGGCTAACACATCCCGAAGATGATTTTGGAGAGATTAATGTTATTTA 1860
Db      1807 ACCAGATTGTTGGCTAACACATCCCGAAGATGATTTTGGAGAGATTAATGTTATTTA 1866
QY      1861 ATAAATATTTCAGGATATTTTCCCTACATTAAGTAACTTAATTAAGTAACTTA 1910
Db      1867 ATAAATATTTCAGGATATTTTCCCTACATTAAGTAACTTAATTAAGTAACTTA 1916

RESULT 5
AF052135      1462 bp      mRNA      linear      PRI 05-AUG-1998
LOCUS      AF052135
DEFINITION      Homo sapiens clone 23625 mRNA sequence.
ACCESSION      AF052135
VERSION      AF052135.1 GI:3360444
KEYWORDS      FLI CDNA.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1462)
AUTHORS      Anderson,B., Wentland,M.A., Ricalfente,J.Y., Liu,W. and Gibbs,R.A.
TITLE      A 'double adaptor' method for improved shotgun library construction
JOURNAL      Anal. Biochem. 236 (1), 107-113 (1996)
MEDLINE      96207227
PUBMED      8619474
REFERENCE      2 (bases 1 to 1462)
AUTHORS      Yu,W., Anderson,B., Morley,K.C., Muzny,D.M., Ding,Y., Liu,W.,
Ricalfente,J.Y., Wentland,M.A., Lennon,G. and Gibbs,R.A.
TITLE      Large-scale concatenation cDNA sequencing
JOURNAL      Genome Res. 7 (4), 353-358 (1997)
MEDLINE      9110174
PUBMED      9110174
REFERENCE      3 (bases 1 to 1462)
AUTHORS      Yu,W., Sarginson,J. and Gibbs,R.A.
TITLE      Submitted (05-MAR-1998) Molecular and Human Genetics, Baylor
JOURNAL      College of Medicine, One Baylor Plaza S930, Houston, TX 77030, USA
FEATURES
     source
         1..1462
             Location/Qualifiers
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="I.M.A.G.E. Consortium clone ID 23625"
             /sex="female"
             /tissue_type="brain"
             /clone_id="INIB"
             /dev_stage="infant"

ORIGIN

Query Match      71.0%; Score 1356; DB 9; Length 1462;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTGGTCTGATGTCGACCATGAGATGAGAGCTCCGCCGGAAGACCGGGTGAAGGC 60
Db      107 CTTGGTCTGATGTCGACCATGAGATGAGAGCTCCGCCGGAAGACCGGGTGAAGGC 166
QY      107 CTTGGTCTGATGTCGACCATGAGATGAGAGCTCCGCCGGAAGACCGGGTGAAGGC 166
Db      167 TCTCTCCAGCTGGTGGTGGTGAAGAGTGAAGATTCACCCCGTGGTACTT 226
QY      121 CCGCTCGAGTGAAGTATCCGAATGCGATCATTTACTTGAAGAGGCAACATTGA 180
Db      227 CCGCTCGAGTGAAGTATCCGAATGCGATCATTTACTTGAAGAGGCAACATTGA 286
QY      181 ACATGCTTATCTCTATATAACAAGTATACGCTCTTTATTGAGAAACTACCAAAACA 240
Db      287 ACATGCTTATCTCTATATAACAAGTATACGCTCTTTATTGAGAAACTACCAAAACA 346
QY      241 TCAGATTACAAATTCGCTGCTCATTCCTGAAAAGAAAGACAGTAAAGAAATTAAAGA 300

```

```

Db      347 TCAGATTACAAATTCGCTGCTCATTCCTGAAAAGAAAGACAGTAAAGAAATTAAAGA 406
QY      301 GATTGCTATTTCCCAAGCAGAAAGCTGAAGGACAGCTGTTAAACGATATACCAAGA 360
Db      407 GATTGCTATTTCCCAAGCAGAAAGCTGAAGGACAGCTGTTAAACGATATACCAAGA 466
QY      361 ATATACGAATATATAGAAAAGAAAGAGAGAGAGAGATTTGGCCCGCAATGCGC 420
Db      467 ATATACGAATATATAGAAAAGAAAGAGAGAGAGATTTGGCCCGCAATGCGC 526
QY      421 CATTCAGCAAGAGCTGAAAAAGAAAAAGAGAGTACACACAGAGACGACGACAT 480
Db      527 CATTCAGCAAGAGCTGAAAAAGAAAAAGAGAGTACACACAGAGACGACGACAT 586
QY      481 GGAACAGGAACAGTCCATGCTTGAAGAGATGATCCGGAAACGAGACTAGAAAAA 540
Db      587 GGAACAGGAACAGTCCATGCTTGAAGAGATGATCCGGAAACGAGAGCTGAAAAA 646
QY      541 GCGACTGAAAATTGTACAGAGATTGGGAAGTAGACCTTGCGCTAGTGCGCCGCTAGT 600
Db      647 GCGACTGAAAATTGTACAGAGATTGGGAAGTAGACCTTGCGCTAGTGCGCCGCTAGT 706
QY      601 GCGTGACTTGGAGAGACCTTCTTATATGTTTCCCACTTAAACATGCTCATCCATAC 660
Db      707 GCGTGACTTGGAGAGACCTTCTTATATGTTTCCCACTTAAACATGCTCATCCATAC 766
QY      661 GCGTTCAGACTGTCACACAACTGTAAAGCCAGCTAAGCACCCTGTGTGAGAGGCTCT 720
Db      767 GCGTTCAGACTGTCACACAACTGTAAAGCCAGCTAAGCACCCTGTGTGAGAGGCTCT 826
QY      721 GAAACTGTGAGACCTGAGCAACTCAGAAAGTATTCACAAATGATGATTCGCCATGT 780
Db      827 GAAACTGTGAGACCTGAGCAACTCAGAAAGTATTCACAAATGATGATTCGCCATGT 886
QY      781 GGTGGTGGCTGGGGGGCTGTGGCCACAGTTTCCAGTAAAGCCATGTCCTAATCTGCGC 840
Db      887 GGTGGTGGCTGGGGGGCTGTGGCCACAGTTTCCAGTAAAGCCATGTCCTAATCTGCGC 946
QY      841 GGGAGTGAAGACATGTGAAATTCCTGTGGAAGAACTGATGAGAAATTAATTAACATTAC 900
Db      947 GGGAGTGAAGACATGTGAAATTCCTGTGGAAGAACTGATGAGAAATTAATTAACATTAC 1006
QY      901 CCAATGTTTCATCCCAAGCAAGAGCTGGCTGATTAATCTGCAACACAGAGAAAGAGA 960
Db      1007 CCAATGTTTCATCCCAAGCAAGAGCTGGCTGATTAATCTGCAACACAGAGAAAGAGA 1066
QY      961 AGAATCTTTTCTCATACAGAGATCAGAGAGGCTCATCACTGGGCTGATTCATATCA 1020
Db      1067 AGAATCTTTTCTCATACAGAGATCAGAGAGGCTCATCACTGGGCTGATTCATATCA 1126
QY      1021 CCCCACACAGACCGGTTTCTCTCAGTGTGACCTTAACACTACTGCTTTACAGAT 1080
Db      1127 CCCCACACAGACCGGTTTCTCTCAGTGTGACCTTAACACTACTGCTTTACAGAT 1186
QY      1081 GATGTTGGCAAGTAGTAGTACATTTGTTGCTCCCGCAAGTTTCCAGGAATCTGATTTCT 1140
Db      1187 GATGTTGGCAAGTAGTAGTACATTTGTTGCTCCCGCAAGTTTCCAGGAATCTGATTTCT 1246
QY      1141 TAACTAACTGACATGAGACTAGAGAGATTTCTCTGTGCGCAAGAAAGATTTCAATCC 1200
Db      1247 TAACTAACTGACATGAGACTAGAGAGATTTCTCTGTGCGCAAGAAAGATTTCAATCC 1306
QY      1201 ACAGAGCAAGAGATCCAGCTCTGTTCTGAGTGCAGACCACTGATGTTGTGGACAGAGC 1260
Db      1307 ACAGAGCAAGAGATCCAGCTCTGTTCTGAGTGCAGACCACTGATGTTGTGGACAGAGC 1366
QY      1261 AGTGAACATCAAGAGCTTGCATGAGAGCTTTGAGTCCAACTCTTCCAAAGCAAAACA 1320
Db      1367 AGTGAACATCAAGAGCTTGCATGAGAGCTTTGAGTCCAACTCTTCCAAAGCAAAACA 1426
QY      1321 CCAATTCAGTGTACTGTAGCCCTTAATTTAAGCTT 1356

```


Best Local Similarity 100.0%; Pred. No. 2.5e-244;
Matches 1043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      863 CTCTGTGAAAACCTGATGAGAAATGATTTCCATTACCCAGTTCTCATCCCAAGCAA 922
Db      1 CTCTGTGAAAACCTGATGAGAAATGATTTCCATTACCCAGTTCTCATCCCAAGCAA 60
QY      923 AGTGTGTGGCTGTGATTAATGTCACACAGAGAACAGAAAGAACTTTCTCATACAGAT 982
Db      61 AGTGTGTGGCTGTGATTAATGTCACACAGAGAACAGAAAGAACTTTCTCATACAGAT 120
QY      983 CAGCAGGCGCTCATACACTGCGGCTGATTCATACACCCACACAGACCGCGTTTCTC 1042
Db      121 CAGCAGGCGCTCATACACTGCGGCTGATTCATACACCCACACAGACCGCGTTTCTC 180
QY      1043 TCCAGTGTGACACTACACTCAGTGGCTTACAGAGATGATGTGCGAGAGTCACTAGCC 1102
Db      181 TCCAGTGTGACACTACACTCAGTGGCTTACAGAGATGATGTGCGAGAGTCACTAGCC 240
QY      1103 ATTGTGTGCTCCCAAGTTCCAGAAACTGATTTCTTTAACTAACTGACCACTGACATA 1162
Db      241 ATTGTGTGCTCCCAAGTTCCAGAAACTGATTTCTTTAACTAACTGACCACTGACATA 300
QY      1163 GAGGAGATTTCTCTGCTGCGCAAAAGATTTTATCCACACAGAAAGATTCACCTCTG 1222
Db      301 GAGGAGATTTCTCTGCTGCGCAAAAGATTTTATCCACACAGAAAGATTCACCTCTG 360
QY      1223 TTCTGTAGCTGACAGCCAGTACTGTGTGGAACAGACAGTACATCACAGACCTTCCA 1282
Db      361 TTCTGTAGCTGACAGCCAGTACTGTGTGGAACAGACAGTACATCACAGACCTTCCA 420
QY      1283 TGAGCGTTTGTAGTCCAAACCTTCCAAAGAACAAACAACTATCACTGACTGTAGCCC 1342
Db      421 TGAGCGTTTGTAGTCCAAACCTTCCAAAGAACAAACAACTATCACTGACTGTAGCCC 480
QY      1343 CTTAATTTAAGCTTTCTAGAAAGCTTTGGAAGTTTGTAGATAGTGAAGGGGGCAT 1402
Db      481 CTTAATTTAAGCTTTCTAGAAAGCTTTGGAAGTTTGTAGATAGTGAAGGGGGCAT 540
QY      1403 CACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGAAAAGAAATTAAGTAAATTTT 1462
Db      541 CACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGAAAAGAAATTAAGTAAATTTT 600
QY      1463 TTAGGCAAGTCAAGAAAGAAACATGTGCACCCAAAAGCAACTGTAACTGAAATTAAGT 1522
Db      601 TTAGGCAAGTCAAGAAAGAAACATGTGCACCCAAAAGCAACTGTAACTGAAATTAAGT 660
QY      1523 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTATATATGAAACCCCATATAC 1582
Db      661 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTATATATGAAACCCCATATAC 720
QY      1583 CCTTCCCTTGTGATTCACCAATTTGTTAACTTTTCTCTCAGCTATCTTTAAATTT 1642
Db      721 CCTTCCCTTGTGATTCACCAATTTGTTAACTTTTCTCTCAGCTATCTTTAAATTT 780
QY      1643 CTCTCTAATTTCAATTTGTTATTTATTTACCTGCGGCTCAATGAAGGCAATCTGTAGAA 1702
Db      781 CTCTCTAATTTCAATTTGTTATTTATTTACCTGCGGCTCAATGAAGGCAATCTGTAGAA 840
QY      1703 ATTTGGAAGCAATTTGAAAAATCTTTTGAATTTCTGTGTGTTATGGAATATGATG 1762
Db      841 ATTTGGAAGCAATTTGAAAAATCTTTTGAATTTCTGTGTGTTATGGAATATGATG 900
QY      1763 AGCTTTAATCTGGGGTGAAGGAGACGCTTACTCATTTGACAGATTTGTTGGTAAACA 1822
Db      901 AGCTTTAATCTGGGGTGAAGGAGACGCTTACTCATTTGACAGATTTGTTGGTAAACA 960
QY      1823 TCCCGAAGATGATTTGTGAGGAATTAATGTTAATTAATTAATTTTCAAGATATTTT 1882
Db      961 TCCCGAAGATGATTTGTGAGGAATTAATGTTAATTAATTAATTTTCAAGATATTTT 1020
QY      1883 CCTCTACATTAAGTAACTAATTA 1905

```

Db 1021 CCTCTACATTAAGTAACTAATTA 1043

RESULT 8
AR257578 1043 bp DNA linear PAT 20-DEC-2002
LOCUS AR257578
DEFINITION Sequence 19 from patent US 6488931.
ACCESSION AR257578
VERSION AR257578.1 GI:27307653
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 1043)
AUTHORS Mitcham, J.L., Frudakis, T.N. and King, G.E.
TITLE Compositions, and methods for therapy and diagnosis of ovarian cancer
JOURNAL Patent: US 6488931-A 19 03-DEC-2002;
FEATURES Location/Qualifiers
source 1..1043
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 54.6%; Score 1043; DB 6; Length 1043;
Best Local Similarity 100.0%; Pred. No. 2.5e-244;
Matches 1043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      863 CTCTGTGAAAACCTGATGAGAAATGATTTACCATTTACCATTTCTCATCCCAAGCAA 922
Db      1 CTCTGTGAAAACCTGATGAGAAATGATTTACCATTTACCATTTCTCATCCCAAGCAA 60
QY      923 AGTGTGTGGCTGTGATTAATGTCACACAGAGAACAGAAAGAACTTTCTCATACAGAT 982
Db      61 AGTGTGTGGCTGTGATTAATGTCACACAGAGAACAGAAAGAACTTTCTCATACAGAT 120
QY      983 CAGCAGGCGCTCATACACTGCGGCTGATTCATACACCCACACAGACCGCGTTTCTC 1042
Db      121 CAGCAGGCGCTCATACACTGCGGCTGATTCATACACCCACACAGACCGCGTTTCTC 180
QY      1043 TCCAGTGTGACACTACACTCAGTGGCTTACAGAGATGATGTGCGAGAGTCACTAGCC 1102
Db      181 TCCAGTGTGACACTACACTCAGTGGCTTACAGAGATGATGTGCGAGAGTCACTAGCC 240
QY      1103 ATTGTGTGCTCCCAAGTTCCAGAAACTGATTTCTTTAACTAACTGACCACTGACATA 1162
Db      241 ATTGTGTGCTCCCAAGTTCCAGAAACTGATTTCTTTAACTAACTGACCACTGACATA 300
QY      1163 GAGGAGATTTCTCTGCTGCGCAAAAGATTTTATCCACACAGAAAGATTCACCTCTG 1222
Db      301 GAGGAGATTTCTCTGCTGCGCAAAAGATTTTATCCACACAGAAAGATTCACCTCTG 360
QY      1223 TTCTGTAGCTGACAGCCAGTACTGTGTGGAACAGACAGTACATCACAGACCTTCCA 1282
Db      361 TTCTGTAGCTGACAGCCAGTACTGTGTGGAACAGACAGTACATCACAGACCTTCCA 420
QY      1283 TGAGCGTTTGTAGTCCAAACCTTCCAAAGAACAAACAACTATCACTGACTGTAGCCC 1342
Db      421 TGAGCGTTTGTAGTCCAAACCTTCCAAAGAACAAACAACTATCACTGACTGTAGCCC 480
QY      1343 CTTAATTTAAGCTTTCTAGAAAGCTTTGGAAGTTTGTAGATAGTGAAGGGGGCAT 1402
Db      481 CTTAATTTAAGCTTTCTAGAAAGCTTTGGAAGTTTGTAGATAGTGAAGGGGGCAT 540
QY      1403 CACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGAAAAGAAATTAAGTAAATTTT 1462
Db      541 CACCTGAGAAAGAGCTGATTTGTATTTCAAGTTTGAAAAGAAATTAAGTAAATTTT 600
QY      1463 TTAGGCAAGTCAAGAAAGAAACATGTGCACCCAAAAGCAACTGTAACTGAAATTAAGT 1522
Db      601 TTAGGCAAGTCAAGAAAGAAACATGTGCACCCAAAAGCAACTGTAACTGAAATTAAGT 660
QY      1523 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTATATGAAACCCCATATAC 1582

```

```

Db      661  TACTCAGAAATTAGTACGTGCAAAATTAGAAAAGATGTAATATGAAACCCCATATAC
Qy      1583  CCTTCCTTGGAATTCACCAATTTTAACTTTTCTCTCAGCTATCCTTCTAATT 1642
Db      721  CCTTCCTTGGAATTCACCAATTTTAACTTTTCTCTCAGCTATCCTTCTAATT 780
Qy      1643  CTTCTAATTTGATTTGTTTATTTATTTACCTCGGGCTCAATAGGGACCTGTGCAGAA 1702
Db      781  CTTCTAATTTGATTTGTTTATTTATTTACCTCGGGCTCAATAGGGACCTGTGCAGAA 840
Qy      1703  ATTGGAAGCATTATGAAAATCTTTTGATTTCTGTGTTTATGCAATATGATGG 1762
Db      841  ATTGGAAGCATTATGAAAATCTTTTGATTTCTGTGTTTATGCAATATGATGG 900
Qy      1763  AGCTTATTTACTGGGGTGAAGGAGAGCTTACTCCATTTGACAGATTTGTGCTAACAA 1822
Db      901  AGCTTATTTACTGGGGTGAAGGAGAGCTTACTCCATTTGACAGATTTGTGCTAACAA 960
Qy      1823  TCCCGAAGATGATTTGTGAGAAATTTATTTATTTAATAATATTTGAGATATTTT 1882
Db      961  TCCCGAAGATGATTTGTGAGAAATTTATTTATTTAATAATATTTGAGATATTTT 1020
Qy      1883  CCTCTACATTAATAGTAACAATTA 1905
Db      1021  CCTCTACATTAATAGTAACAATTA 1043

RESULT 9
LOCUS      AR283624      1043 bp      DNA      linear      PAT 10-APR-2003
DEFINITION Sequence 19 from patent US 6528253.
ACCESSION  AR283624
VERSION     AR283624.1  GI:29720521
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1043)
AUTHORS    Mitcham,J.L., Frudakis,T.N. and King,G.E.
TITLE      Compositions and methods for diagnosis of ovarian cancer
JOURNAL    Patent: US 6528253-A 19 04-MAR-2003;
FEATURES
SOURCE      Location/Qualifiers
             1..1043
             /organism="unknown"
             /mol_type="genomic DNA"

ORIGIN
Query Match      54.6%; Score 1043; DB 6; Length 1043;
Best Local Similarity 100.0%; Pred. No. 2, 5e-244;
Matches 1043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      863  CTCTGTGAAAACCTGATGAGATGAATTTACCATTAACCATGTTCTCATCCCAAGCAA 922
Db      1  CTCTGTGAAAACCTGATGAGATGAATTTACCATTAACCATGTTCTCATCCCAAGCAA 60
Qy      923  AGTGTCTGGTCTGATTTCTGTGCAACAGAGAGAGAAAGAACTTTTCTCATACAGAT 982
Db      61  AGTGTCTGGTCTGATTTCTGTGCAACAGAGAGAGAAAGAACTTTTCTCATACAGAT 120
Qy      983  CAGCAGGGCCTCATACACTGAGCTGATTCATCTACCCCAACACAGACCGGTTTCTC 1042
Db      121  CAGCAGGGCCTCATACACTGAGCTGATTCATCTACCCCAACACAGACCGGTTTCTC 180
Qy      1043  TCCAGTGTGACCTTACACACTCATCTGCTTTTACCAAGATGTTGTCAGAGTCAGTACC 1102
Db      181  TCCAGTGTGACCTTACACACTCATCTGCTTTTACCAAGATGTTGTCAGAGTCAGTACC 240
Qy      1103  ATTGTTGCTCCCAAGTTCAGAGAACTGATTTCTTTAACTTACGACATGACATA 1162
Db      241  ATTGTTGCTCCCAAGTTCAGAGAACTGATTTCTTTAACTTACGACATGACATA 300
Qy      1163  GAGGAGATTTCTCTCTGTGCGCAAGAAAGATTTATCCACACAGCAAGATCCACTCTG 1222

```

```

Db      301  GAGGAGATTTCTCTGTGCGCAAGAAAGATTTATCCACACAGCAAGATCCACTCTG 360
Qy      1223  TTCTGTAGCTGACGCCAGTGAATTTGTGGAAGAGTGAACCATTCACAGACTTCGA 1282
Db      361  TTCTGTAGCTGACGCCAGTGAATTTGTGGAAGAGTGAACCATTCACAGACTTCGA 420
Qy      1283  TGAAGCTTTGATGTCACACCTTCCAAAGAACCAAAACCATATTCAGTACTGTAGCCC 1342
Db      421  TGAAGCTTTGATGTCACACCTTCCAAAGAACCAAAACCATATTCAGTACTGTAGCCC 480
Qy      1343  CTTAATTTAGCTTTCTAGAAAGCTTTGGAAGTTTGTGATATGATGAAAGGGGGCAT 1402
Db      481  CTTAATTTAGCTTTCTAGAAAGCTTTGGAAGTTTGTGATATGATGAAAGGGGGCAT 540
Qy      1403  CACCTGAGAAAGACCTGATTTTGTATTTGAGTTTGAAAGAAATTAATCTAATATTT 1462
Db      541  CACCTGAGAAAGACCTGATTTTGTATTTGAGTTTGAAAGAAATTAATCTAATATTT 600
Qy      1463  TTAGGCAAGTCAGAAAGAACATGTCACCCCAAGCACTGTAATCTGAAATTAAGT 1522
Db      601  TTAGGCAAGTCAGAAAGAACATGTCACCCCAAGCACTGTAATCTGAAATTAAGT 660
Qy      1523  TACTCAGAAATTAAGTACCTCAGAAATTAAGAAAGATGTTATATGAAACCCCATATAC 1582
Db      661  TACTCAGAAATTAAGTACCTCAGAAATTAAGAAAGATGTTATATGAAACCCCATATAC 720
Qy      1583  CCTTCCTTGGAATTCACCAATTTTGTAACTTTTCTCTCAGCTATCCTTCTAATT 1642
Db      721  CCTTCCTTGGAATTCACCAATTTTGTAACTTTTCTCTCAGCTATCCTTCTAATT 780
Qy      1643  CTTCTAATTTCAATTTGTTTATTTATTTACCTCGGGCTCAATAGGGCATCTGTGCAGAA 1702
Db      781  CTTCTAATTTCAATTTGTTTATTTATTTACCTCGGGCTCAATAGGGCATCTGTGCAGAA 840
Qy      1703  ATTGGAAGCATTATGAAAATCTTTTGATTTCTGTGTTTATGCAATATGATGG 1762
Db      841  ATTGGAAGCATTATGAAAATCTTTTGATTTCTGTGTTTATGCAATATGATGG 900
Qy      1763  AGCTTATTTACTGGGGTGAAGGAGAGCTTACTCCATTTGACAGATTTGTGCTAACAA 1822
Db      901  AGCTTATTTACTGGGGTGAAGGAGAGCTTACTCCATTTGACAGATTTGTGCTAACAA 960
Qy      1823  TCCCGAAGATGATTTGTGAGAAATTTATTTATTTAATAATATTTGAGATATTTT 1882
Db      961  TCCCGAAGATGATTTGTGAGAAATTTATTTATTTAATAATATTTGAGATATTTT 1020
Qy      1883  CCTCTACATTAATAGTAACAATTA 1905
Db      1021  CCTCTACATTAATAGTAACAATTA 1043

RESULT 10
LOCUS      AX366252      1043 bp      DNA      linear      PAT 15-FEB-2002
DEFINITION Sequence 19 from Patent W00206317.
ACCESSION  AX366252
VERSION     AX366252.1  GI:18697678
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS    Mitcham,J.L., King,G.E., Algate,P.A., Fling,S.P., Retter,M.W.,
           Fanger,G.R., Reed,S.G., Vedvick,T.S., Carter,D., Hill,P. and
           Albone,B.
TITLE      Compositions and methods for the therapy and diagnosis of ovarian
           cancer
JOURNAL    Patent: WO 0206317-A 19 24-JAN-2002;
           CORIXA CORPORATION (US)
FEATURES
SOURCE      Location/Qualifiers
             1..1043

```

ORIGIN

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 54.6%; Score 1043; DB 6; Length 1043;
Best Local Similarity 100.0%; Pred. No. 2,56-244; Indels 0; Gaps 0;
Matches 1043; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 863 CTCTGTGAAAAAAGTATGAGATGATATTTACCATTCACCATGTTTCATCCCAAGCAA 922
DB 1 CTCTGTGAAAAAAGTATGAGATGATATTTACCATTCACCATGTTTCATCCCAAGCAA 60
QY 923 AGTGTGGGTCTGATTATCGCAACACAGAAACGAAAGAACTTTCCCTCATACAGAT 982
DB 61 AGTGTGGGTCTGATTATCGCAACACAGAAACGAAAGAACTTTCCCTCATACAGAT 120
QY 983 CAGCAGGGCTCATCACTGGGCTGATATCTACCCACACAGACCGCTTCTC 1042
DB 121 CAGCAGGGCTCATCACTGGGCTGATATCTACCCACACAGACCGCTTCTC 180
QY 1043 TCCAGTGTGACCTACACACTCACTGCTTTACCAAGATGATGTTGCCAGATCGTACC 1102
DB 181 TCCAGTGTGACCTACACACTCACTGCTTTACCAAGATGATGTTGCCAGATCGTACC 240
QY 1103 ATTGTTGCTCCCAAGTTCCAGAAATGATTTCTTTAACTACTGACCATGACTA 1162
DB 241 ATTGTTGCTCCCAAGTTCCAGAAATGATTTCTTTAACTACTGACCATGACTA 300
QY 1163 GAGAGATTTTCTCTGTGCGCAAGAAAGATTTATCCACACAGCAAGATCCACTCTG 1222
DB 301 GAGAGATTTTCTCTGTGCGCAAGAAAGATTTATCCACACAGCAAGATCCACTCTG 360
QY 1223 TTCTGTAGCTCAGCCAGTCACTGTTGTGGACAGACAGTGCATCAAGACCTTCCA 1282
DB 361 TTCTGTAGCTCAGCCAGTCACTGTTGTGGACAGACAGTGCATCAAGACCTTCCA 420
QY 1283 TGAGCGTTTGAAGTCAACACCTTCCAGAAACAAACCAATACATGATGATGACCC 1342
DB 421 TGAGCGTTTGAAGTCAACACCTTCCAGAAACAAACCAATACATGATGATGACCC 480
QY 1343 CTTAATTTAAGCTTTTGAAGAGCTTTGGAAGTTTGTGATGATGAAAGGGGGCAT 1402
DB 481 CTTAATTTAAGCTTTTGAAGAGCTTTGGAAGTTTGTGATGATGAAAGGGGGCAT 540
QY 1403 CACCTGAGAAAGAGCTGATTTTGTATTTCAAGTTTGAAGAAATTAAGTAACTAATTT 1462
DB 541 CACCTGAGAAAGAGCTGATTTTGTATTTCAAGTTTGAAGAAATTAAGTAACTAATTT 600
QY 1463 TTATGAGCAAGTCAAGAAAGAAACATGTCACCCAAAGCACTGTAACTGAAATTAAGT 1522
DB 601 TTATGAGCAAGTCAAGAAAGAAACATGTCACCCAAAGCACTGTAACTGAAATTAAGT 660
QY 1523 TACTCAAAATTAAGTCAAGAAATTAAGAAAGATGATTAAGTAAAGCCCATATAC 1582
DB 661 TACTCAAAATTAAGTCAAGAAATTAAGAAAGATGATTAAGTAAAGCCCATATAC 720
QY 1583 CCTTCTCTGATCAACCAATTTTGAATTTTCTCTCAAGTATCTCTTAATTT 1642
DB 721 CCTTCTCTGATCAACCAATTTTGAATTTTCTCTCAAGTATCTCTTAATTT 780
QY 1643 CTCTGTAATTTCAATTTGTTTATTTAATTTAATTTCTGGGCTCAATAGGGCATCTGTGAGAA 1702
DB 781 CTCTGTAATTTCAATTTGTTTATTTAATTTAATTTCTGGGCTCAATAGGGCATCTGTGAGAA 840
QY 1703 ATTGGAAGCATTTAGAAATCTTTTGAATTTTCTGTGTTATGGAATATGATG 1762
DB 841 ATTGGAAGCATTTAGAAATCTTTTGAATTTTCTGTGTTATGGAATATGATG 900
QY 1763 AGCTTTTAACTGGGGTGAAGGACAGCTTACTCATTTGACAGATTTGTTGGCTAACACA 1822
DB 901 AGCTTTTAACTGGGGTGAAGGACAGCTTACTCATTTGACAGATTTGTTGGCTAACACA 960

```

QY 1823 TCCGAAAGATGATTTTGTGAGAAATTAATTTAATTAATTAATTTTCAAGATATTTT 1882
DB 961 TCCGAAAGATGATTTTGTGAGAAATTAATTTAATTAATTAATTTTCAAGATATTTT 1020
QY 1883 CCTTCAATTAAGTAACTAATTTA 1905
DB 1021 CCTTCAATTAAGTAACTAATTTA 1043

RESULT 11
BD264633
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

OS Homo sapiens (human)
PN JP 2002532093-A/18
PD 02-OCT-2002
PE 17-DEC-1999 JP 2000588356
PR 17-DEC-1998 US 09/215681, 17-DEC-1998 US 09/215603 PR
23-JUN-1999 US 09/338933, 24-SEP-1999 US 09/404879 FI
JENNIFER L MITCHEM, GORDON E KING, PAUL A ALGATE, TONY N FRUDAKIS PC
C12N15/09, A61K31/7115, A61K35/14, A61K35/76, A61K39/00, A61K39/395, PC
A61K48/00, A61P35/00, A61P37/04, C07K14/82, C07K19/00, C12N1/15, PC
C12N1/19, PC

PC C12N1/21, C12N5/06, C12N5/10, C12N1/68, G01N33/53, G01N33/53, G01N33/ PC
566, G01N33/574, G01N33/577, C12N15/00, C12N5/00, C12N5/00 CC n =
PC A,T,C or G Location/Qualifiers
FH Key (1) . (1041) .
FT misc feature (1) . (1041) .

FEATURES
source
1..1041
Location/Qualifiers

1..1041
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 54.5%; Score 1040; DB 6; Length 1041;
Best Local Similarity 99.9%; Pred. No. 1,46-243; Indels 1; Gaps 0;
Matches 1040; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

```

QY 863 CTCTGTGAAAAAAGTATGAGATGATATTTACCATTCACCATGTTTCATCCCAAGCAA 922
DB 1 CTCTGTGAAAAAAGTATGAGATGATATTTACCATTCACCATGTTTCATCCCAAGCAA 60
QY 923 AGTGTGGGTCTGATTATCGCAACACAGAAACGAAAGAACTTTCCCTCATACAGAT 982
DB 61 AGTGTGGGTCTGATTATCGCAACACAGAAACGAAAGAACTTTCCCTCATACAGAT 120
QY 983 CAGCAGGGCTCATCACTGGGCTGATATCTACCCACACAGACCGCTTCTC 1042
DB 121 CAGCAGGGCTCATCACTGGGCTGATATCTACCCACACAGACCGCTTCTC 180
QY 1043 TCCAGTGTGACCTACACACTCACTGCTTTACCAAGATGATGTTGCCAGATCGTACC 1102
DB 181 TCCAGTGTGACCTACACACTCACTGCTTTACCAAGATGATGTTGCCAGATCGTACC 240
QY 1103 ATTGTTGCTCCCAAGTTCCAGAAATGATTTCTTTAACTACTGACCATGACTA 1162

```


Db 241 ATTGTTGCTCCCAAGTTCAGAGAACTGATTTCTTAATACTGACGACATGACTA 300
 QY 1163 GAGGAGATTTCTTCTGTGCGCAAGAGATTTCACTCCACAGAGAGATCCACTCTG 1222
 Db 301 GAGGAGATTTCTTCTGTGCGCAAGAGATTTCACTCCACAGAGAGATCCACTCTG 360
 QY 1223 TTCTGTAGTGCAGCAGCTGACTGTTGTGACAGAGAGAGATCCAGACTCTGCA 1282
 Db 361 TTCTGTAGTGCAGCAGCTGACTGTTGTGACAGAGAGAGATCCAGACTCTGCA 420
 QY 1283 TGAGCGTTGAGTCCACACCTTCCAGAGACAAACCAATACAGTGTACTGACCC 1342
 Db 421 TGAGCGTTGAGTCCACACCTTCCAGAGACAAACCAATACAGTGTACTGACCC 480
 QY 1343 CTTAATTTAGCTTTCTGAGAAAGCTTTGAGATTTGTAGATAGTGAAGAGGCGCAT 1402
 Db 481 CTTAATTTAGCTTTCTGAGAAAGCTTTGAGATTTGTAGATAGTGAAGAGGCGCAT 540
 QY 1403 CACCTGAGAAAGAGCTGATTTGTATTTCAAGTTGAAAAGAAATTAATGACATATTT 1462
 Db 541 CACNTGAGAAAGAGCTGATTTGTATTTCAAGTTGAAAAGAAATTAATGACATATTT 600
 QY 1463 TTAGCGAAGTCAGAAAGAGACATGTCACCCAAAGCACTGTAACCTGAGAAATTAAGT 1522
 Db 601 TTAGCGAAGTCAGAAAGAGACATGTCACCCAAAGCACTGTAACCTGAGAAATTAAGT 660
 QY 1523 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTAATATGAAACCCCATATAC 1582
 Db 661 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTAATATGAAACCCCATATAC 720
 QY 1583 CCTTCCTCTGATTCACCAATGTTGTAATTTTCTCTGATCTGATCTCTGATCTCTGAT 1642
 Db 721 CCTTCCTCTGATTCACCAATGTTGTAATTTTCTCTGATCTGATCTCTGATCTCTGAT 780
 QY 1643 CTCTCTAATTTAATTTGTTATTTATTTAATTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 1702
 Db 781 CTCTCTAATTTAATTTGTTATTTATTTAATTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 840
 QY 1703 ATTTGGAAGCAATTAAGAAATCTTTTGAATTTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 1762
 Db 841 ATTTGGAAGCAATTAAGAAATCTTTTGAATTTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 900
 QY 1763 AGCTTATTAATCTGAGGTCAGGAGACAGCTTACCTCATTGACCAAGATTTGTCATACACA 1822
 Db 901 AGCTTATTAATCTGAGGTCAGGAGACAGCTTACCTCATTGACCAAGATTTGTCATACACA 960
 QY 1823 TCCCGAAGATGATTTTGTGAGAAATTAATTTAATTAATTAATTTGAGATTTT 1882
 Db 961 TCCCGAAGATGATTTTGTGAGAAATTAATTTAATTAATTAATTTGAGATTTT 1020
 QY 1883 CCTCTACATTAAGTAACAT 1903
 Db 1021 CCTCTACATTAAGTAACAT 1041

RESULT 12

AR238036

LOCUS AR238036 1041 bp DNA linear PART 20-DEC-2002

DEFINITION Sequence 18 from patent US 6468546.

ACCESSION AR238036 GI:27283000

VERSION AR238036.1

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

AUTHORS 1 (bases 1 to 1041)

TITLES

JOURNAL

FEATURES

Location/Qualifiers

1..1041

source

Patent: US 6468546-A 18 22-OCT-2002;
 Location/Qualifiers
 1..1041
 /organism="Unknown"

ORIGIN

/mol_type="genomic DNA"

Query Match 54.5%; Score 1040; DB 6; Length 1041;

Best Local Similarity 99.9%; Pred. No. 1,4e-243;

Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 863 CTCGTGGGAAACGATGAGAGATGATTTACCATTTACCATTTCTGATCCCAAGCA 922
 Db 1 CTCGTGGGAAACGATGAGAGATGATTTACCATTTACCATTTCTGATCCCAAGCA 60
 QY 923 AGTGTGGGTCGATTTAATCTGCAACAGAGACGAGAGAACTTTTCTCATACAGAT 982
 Db 61 AGTGTGGGTCGATTTAATCTGCAACAGAGACGAGAGAACTTTTCTCATACAGAT 120
 QY 983 CAGAGGCGCTCATACAGCTGGGCTGATTTCAATACACCCAGACAGACCGGTTCTC 1042
 Db 121 CAGAGGCGCTCATACAGCTGGGCTGATTTCAATACACCCAGACAGACCGGTTCTC 180
 QY 1043 TCCAGTGTGACCTTACCACTGCTCTTACCAAGTGAATGTTGCCAGATGATAGCC 1102
 Db 181 TCCAGTGTGACCTTACCACTGCTCTTACCAAGTGAATGTTGCCAGATGATAGCC 240
 QY 1103 ATTGTTGCTCCCAAGTTCAGAGAACTGATTTCTTTAACTAATGACATGAGACTA 1162
 Db 241 ATTGTTGCTCCCAAGTTCAGAGAACTGATTTCTTTAACTAATGACATGAGACTA 300
 QY 1163 GAGAGATTTCTTCTGTGCGCAAGAGATTTCAATCCACAGAGAGATCCACTCTG 1222
 Db 301 GAGAGATTTCTTCTGTGCGCAAGAGATTTCAATCCACAGAGAGATCCACTCTG 360
 QY 1223 TTCTGTAGCTGACGACGAGTGTGAGACAGAGAGATGACATCCAGACCTTGA 1282
 Db 361 TTCTGTAGCTGACGACGAGTGTGAGACAGAGAGATGACATCCAGACCTTGA 420
 QY 1283 TGAGCGTTGAGTCCACACCTTCCAGAGACAAACCAATTAATGACATATTT 1342
 Db 421 TGAGCGTTGAGTCCACACCTTCCAGAGACAAACCAATTAATGACATATTT 480
 QY 1343 CTTAATTTAGCTTTCTGAGAAAGCTTTGAGATTTGTAGATAGTGAAGAGGCGCAT 1402
 Db 481 CTTAATTTAGCTTTCTGAGAAAGCTTTGAGATTTGTAGATAGTGAAGAGGCGCAT 540
 QY 1403 CACCTGAGAAAGAGCTGATTTGTATTTCAAGTTGAAAAGAAATTAATGACATATTT 1462
 Db 541 CACNTGAGAAAGAGCTGATTTGTATTTCAAGTTGAAAAGAAATTAATGACATATTT 600
 QY 1463 TTAGCGAAGTCAGAAAGAGACATGTCACCCAAAGCACTGTAACCTGAGAAATTAAGT 1522
 Db 601 TTAGCGAAGTCAGAAAGAGACATGTCACCCAAAGCACTGTAACCTGAGAAATTAAGT 660
 QY 1523 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTAATATGAAACCCCATATAC 1582
 Db 661 TACTCAGAAATTAAGTACTCAGAAATTAAGAAAGATGTAATATGAAACCCCATATAC 720
 QY 1583 CCTTCCTCTGATTCACCAATGTTGTAATTTTCTCTGATCTGATCTCTGATCTCTGAT 1642
 Db 721 CCTTCCTCTGATTCACCAATGTTGTAATTTTCTCTGATCTGATCTCTGATCTCTGAT 780
 QY 1643 CTCTCTAATTTAATTTGTTATTTATTTAATTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 1702
 Db 781 CTCTCTAATTTAATTTGTTATTTATTTAATTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 840
 QY 1703 ATTTGGAAGCAATTAAGAAATCTTTTGAATTTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 1762
 Db 841 ATTTGGAAGCAATTAAGAAATCTTTTGAATTTTCTCTGAGCTCAATTAAGGCAATGTCAGAA 900
 QY 1763 AGCTTATTAATCTGAGGTCAGGAGACAGCTTACCTCATTGACCAAGATTTGTCATACACA 1822
 Db 901 AGCTTATTAATCTGAGGTCAGGAGACAGCTTACCTCATTGACCAAGATTTGTCATACACA 960
 QY 1823 TCCCGAAGATGATTTTGTGAGAAATTAATTTAATTAATTAATTTGAGATTTT 1882

Db 961 TCCGAGAGATGATTTTGTCAAGATTAATTTATTAATAATTTTCAGAGATTTT 1020

Qy 1883 CCTCTACAAATAAGTAACAAT 1903

Db 1021 CCTCTACAAATAAGTAACAAT 1041

RESULT 13

AR257577 1041 bp DNA linear PAT 20-DEC-2002

LOCUS Sequence 18 from patent US 6488931.

DEFINITION AR257577

ACCESSION AR257577

VERSION AR257577.1 GI:27307652

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1041)

AUTHORS Mitcham,J.L., Frudakis,T.N. and King,G.E.

TITLE Compositions and methods for therapy and diagnosis of ovarian cancer

JOURNAL Patent: US 6488931-A 18 03-DEC-2002;

FEATURES

source 1..1041

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 54.5%; Score 1040; DB 6; Length 1041;

Best Local Similarity 99.9%; Pred. No. 1.4e-243;

Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 863 CTCGTGAGAAAATGATGAGAAATGAATTAACCATTAACCATGTTCTCATCCCCAGCA 922

Db 1 CTCGTGAGAAAATGATGAGAAATGAATTAACCATTAACCATGTTCTCATCCCCAGCA 60

Qy 923 AGTGCTGGGTCTGATTAATCTGCAACACAGAGAACGAAAGAACTTTCCCTCATACAGAT 982

Db 61 AGTGCTGGGTCTGATTAATCTGCAACACAGAGAACGAAAGAACTTTCCCTCATACAGAT 120

Qy 983 CAGCAGGGCTCTCATCACTGCGCTGATTCATCTCAACCCACACAGACCGCGTTTCTC 1042

Db 121 CAGCAGGGCTCTCATCACTGCGCTGATTCATCTCAACCCACACAGACCGCGTTTCTC 180

Qy 1043 TCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTGGCAGAGTCAAGACC 1102

Db 181 TCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTGGCAGAGTCAAGACC 240

Qy 1103 ATTGTTTGCTCCCCCAAGTTCCAGGAACTGGAATCTTTAACTAACTGACATGAGACTA 1162

Db 241 ATTGTTTGCTCCCCCAAGTTCCAGGAACTGGAATCTTTAACTAACTGACATGAGACTA 300

Qy 1163 GAGGAGATTTCTCTCCCTGCGCGAAGAAAGATTTTCATCCACACAGCAAGATTCACCTCTG 1222

Db 301 GAGGAGATTTCTCTCCCTGCGCGAAGAAAGATTTTCATCCACACAGCAAGATTCACCTCTG 360

Qy 1223 TTCTGTAGTGCAGCCAGTGAATGTTGTGAGACAGAGCACTGACATCAAGACCTTCGA 1282

Db 361 TTCTGTAGTGCAGCCAGTGAATGTTGTGAGACAGAGCACTGACATCAAGACCTTCGA 420

Qy 1283 TTAGCGCTTAAATGCCAACACCTTTCAGAAACAACAAACATATCAATGTAATGAGCC 1342

Db 421 TTAGCGCTTAAATGCCAACACCTTTCAGAAACAACAAACATATCAATGTAATGAGCC 480

Qy 1343 CTTAATTTAAGCTTTCTAGAAAGCTTTTGAAGTTTGTGATAGATGAAGAAAGGGGGCAT 1402

Db 481 CTTAATTTAAGCTTTCTAGAAAGCTTTTGAAGTTTGTGATAGATGAAGAAAGGGGGCAT 540

Qy 1403 CACCTGAGAAAGAGCTGATTTTGTATTTCAAGTTTGAAGAAATAAATGAACATATTT 1462

Db 541 CACCTGAGAAAGAGCTGATTTTGTATTTCAAGTTTGAAGAAATAAATGAACATATTT 600

Qy 1463 TTAGGCAAGTCAGAAAGAAACATGGTCACCCAAAAGCAACTGTAACTGAGAAATTAAGT 1522

Db 601 TTAGCAGATCGAGAAAGAACATGCTACCCCAAGCACTGTAATCAGAAATTAAGT 660

Qy 1523 TACTCAGAAATTAAGTCTCAGAAATTAAGAAAGATGATTAAGAACCCCATATATC 1582

Db 661 TACTCAGAAATTAAGTCTCAGAAATTAAGAAAGATGATTAAGAACCCCATATATC 720

Qy 1583 CCTTCCTCTGATTCACCAATGTTTACATTTTTCCTCTCAGCTATCTCTAATTT 1642

Db 721 CCTTCCTCTGATTCACCAATGTTTACATTTTTCCTCTCAGCTATCTCTAATTT 780

Qy 1643 CTCTCAATTTCAATTTGTTTATTTATTTACCTCGGGCTCAATTAAGGCATCTGTGAGAA 1702

Db 781 CTCTCAATTTCAATTTGTTTATTTATTTACCTCGGGCTCAATTAAGGCATCTGTGAGAA 840

Qy 1703 ATTGGAAGCAATTTGAGAAATCTTTGGATTTTCTGTGTTTATGCAATATGAAATG 1762

Db 841 ATTGGAAGCAATTTGAGAAATCTTTGGATTTTCTGTGTTTATGCAATATGAAATG 900

Qy 1763 AGCTTAATTAATCTGGGTGAGGAGACAGTTTACCTCAATTTGACAGATTTGTGCTAACA 1822

Db 901 AGCTTAATTAATCTGGGTGAGGAGACAGTTTACCTCAATTTGACAGATTTGTGCTAACA 960

Qy 1823 TCCGAGAGATGATTTGTGAGAAATTAATTTATTTAATAATTTTACAGATATTTT 1882

Db 961 TCCGAGAGATGATTTGTGAGAAATTAATTTATTTAATAATTTTACAGATATTTT 1020

Qy 1883 CCTCTACAAATAAGTAACAAT 1903

Db 1021 CCTCTACAAATAAGTAACAAT 1041

RESULT 14

AR283623 1041 bp DNA linear PAT 10-APR-2003

LOCUS AR283623

DEFINITION Sequence 18 from patent US 6528253.

ACCESSION AR283623

VERSION AR283623.1 GI:29720520

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1041)

AUTHORS Mitcham,J.L., Frudakis,T.N. and King,G.E.

TITLE Compositions and methods for diagnosis of ovarian cancer

JOURNAL Patent: US 6528253-A 18 04-MAR-2003;

FEATURES

source 1..1041

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 54.5%; Score 1040; DB 6; Length 1041;

Best Local Similarity 99.9%; Pred. No. 1.4e-243;

Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 863 CTCGTGAGAAAATGATGAGAAATGAATTAACCATTAACCATGTTCTCATCCCCAGCA 922

Db 1 CTCGTGAGAAAATGATGAGAAATGAATTAACCATTAACCATGTTCTCATCCCCAGCA 60

Qy 923 AGTGCTGGGTCTGATTAATCTGCAACACAGAGAACGAAAGAACTTTCCCTCATACAGAT 982

Db 61 AGTGCTGGGTCTGATTAATCTGCAACACAGAGAACGAAAGAACTTTCCCTCATACAGAT 120

Qy 983 CAGCAGGGCTCTCATCACTGCGCTGATTCATCTCAACCCACACAGACCGCGTTTCTC 1042

Db 121 CAGCAGGGCTCTCATCACTGCGCTGATTCATCTCAACCCACACAGACCGCGTTTCTC 180

Qy 1043 TCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTGGCAGAGTCAAGACC 1102

Db 181 TCCAGTGTGACCTACACACTGCTCTTACCAAGATGATGTGGCAGAGTCAAGACC 240

Qy 1103 ATTGTTTGCTCCCCCAAGTTCCAGGAACTGGAATCTTTAACTAACTGACATGAGACTA 1162

|||||
Db 241 ATGTTGCTCCCCAAGTCCAGAAACGATGATTTTAACTAATCAATGAGACTA 300
1163 GAGAGATTTCTCCGTGCGCAGAAAGATTTTCATCCACAGCAAGATCCACTCTG 1222
Db 301 GAGAGATTTCTCCGTGCGCAGAAAGATTTTCATCCACAGCAAGATCCACTCTG 360
Qy 1223 TTCTGTAGTGCAGCCACGTGACTGTGTGAGCAGAGCAGTACATCAACACTTGC 1282
Db 361 TTCTGTAGTGCAGCCACGTGACTGTGTGAGCAGAGCAGTACATCAACACTTGC 420
Qy 1283 TGAGCGTTTGAATCCACACCTTCCAAAGAACATCTGATGATCTGATGAGCC 1342
Db 421 TGAGCGTTTGAATCCACACCTTCCAAAGAACATCTGATGATCTGATGAGCC 480
Qy 1343 CTAAATTTAAGCTTTCTAGAAAGCTTTGAAATTTTGTAGATGAGAAAGGGGGCAT 1402
Db 481 CTAAATTTAAGCTTTCTAGAAAGCTTTGAAATTTTGTAGATGAGAAAGGGGGCAT 540
Qy 1403 CACCTGAGAAAGAGCTGATTTTGTATTTTCAGTTTGAAGAAATACTGAAATTTT 1462
Db 541 CACCTGAGAAAGAGCTGATTTTGTATTTTCAGTTTGAAGAAATACTGAAATTTT 600
Qy 1463 TTAGGCAAGTGCAGAAAGAGAACTGTCAACCCAAAGCACTGTAATGAAATTTAGT 1522
Db 601 TTAGGCAAGTGCAGAAAGAGAACTGTCAACCCAAAGCACTGTAATGAAATTTAGT 660
Qy 1523 TACTCAGAAATTAAGTACTGAGAAATTAAGAAAGAAATGATTAAGAACCCCATATAC 1582
Db 661 TACTCAGAAATTAAGTACTGAGAAATTAAGAAAGAAATGATTAAGAACCCCATATAC 720
Qy 1583 CCTTCCTCTGAGATCCCAATTTGTTACATTTTCTCAGGATCTCTAATTT 1642
Db 721 CCTTCCTCTGAGATCCCAATTTGTTACATTTTCTCAGGATCTCTAATTT 780
Qy 1643 CTCTCAATTTCAATTTGTTTAAATTTTACCTCTGGGCTCAATAAGGGCATCTGTGAGAA 1702
Db 781 CTCTCAATTTCAATTTGTTTAAATTTTACCTCTGGGCTCAATAAGGGCATCTGTGAGAA 840
Qy 1703 ATTGGAAGCATTTAGAAATCTTTTGATTTTCCGTGGTTTAAAGGCAATTAAGATG 1762
Db 841 ATTGGAAGCATTTAGAAATCTTTTGATTTTCCGTGGTTTAAAGGCAATTAAGATG 900
Qy 1763 AGCTTAATTTACTGGGGTGAGGAGCAGCTTACATTTGACAGATTTGTTGGCTAACACA 1822
Db 901 AGCTTAATTTACTGGGGTGAGGAGCAGCTTACATTTGACAGATTTGTTGGCTAACACA 960
Qy 1823 TCCCGAAGATGATTTTGTGTCAGAAATTTGTTTAAATTAATTTTTCAGATTTT 1882
Db 961 TCCCGAAGATGATTTTGTGTCAGAAATTTGTTTAAATTAATTTTTCAGATTTT 1020
Qy 1883 CCTCTAATTAAGTAAATTAATTAAT 1903
Db 1021 CCTCTAATTAAGTAAATTAATTAAT 1041

RESULT 15
AX366251 1041 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 18 from Patent WO0206317.
DEFINITION AX366251
ACCESSION AX366251 GI:18697677
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Mitcham, J.L., King, G.E., Algate, P.A., Fling, S.P., Retter, M.W.,
Fanger, G.R., Reed, S.G., Vedvick, T.S., Carter, D., Hill, P., and
Albione, E.
Compositions and methods for the therapy and diagnosis of ovarian
cancer

JOURNAL Parent: WO 0206317-A 18 24-JAN-2002;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1041
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 54.5%; Score 1040; DB 6; Length 1041;
Best Local Similarity 99.9%; Pred. No. 1,4e-243;
Matches 1040; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
863 CTCTGTGAAATCGATGAGAAATGATTAACATTAACCATGTTCTCATCCCAAGCA 922
Db 1 CTCTGTGAAATCGATGAGAAATGATTAACATTAACCATGTTCTCATCCCAAGCA 60
Qy 923 AGTGTGGGTCTGATTAATCTGCAACACAGAGACCAAGAACTTTTCTCATCAGAT 982
Db 61 AGTGTGGGTCTGATTAATCTGCAACACAGAGACCAAGAACTTTTCTCATCAGAT 120
Qy 983 CAGAGAGGCTCATACACCTGGGCTGATTAATCTGCAACACAGAGACCGGCTTCTC 1042
Db 121 CAGAGAGGCTCATACACCTGGGCTGATTAATCTGCAACACAGAGACCGGCTTCTC 180
Qy 1043 TCCAGTGTGCACTCAACACTCTGCTTACCAAGATGATTTGCCAGATCAGTACGC 1102
Db 181 TCCAGTGTGCACTCAACACTCTGCTTACCAAGATGATTTGCCAGATCAGTACGC 240
Qy 1103 ATTGTTGCTCCCCAAGTTCAGAAATCTGATTTCTTAACTAATGACCATGAGCTA 1162
Db 241 ATTGTTGCTCCCCAAGTTCAGAAATCTGATTTCTTAACTAATGACCATGAGCTA 300
Qy 1163 GAGAGATTTTCTCTGTCGCGCAGAAAGATTTTATCACAACAGAAAGATTCACCTCTG 1222
Db 301 GAGAGATTTTCTCTGTCGCGCAGAAAGATTTTATCACAACAGAAAGATTCACCTCTG 360
Qy 1223 TTCTGTAGCTGAGCAGATGATCTGTGTGAGCAGAGAGAGACATCAAGACCTTGC 1282
Db 361 TTCTGTAGCTGAGCAGATGATCTGTGTGAGCAGAGAGAGACATCAAGACCTTGC 420
Qy 1283 TGAGCGTTTGAATCCACACCTTCCAAAGAACATCTGATGATCTGATGAGCC 1342
Db 421 TGAGCGTTTGAATCCACACCTTCCAAAGAACATCTGATGATCTGATGAGCC 480
Qy 1343 CTAAATTTAAGCTTTCTAGAAAGCTTTGAAATTTTGTAGATGAGAAAGGGGGCAT 1402
Db 481 CTAAATTTAAGCTTTCTAGAAAGCTTTGAAATTTTGTAGATGAGAAAGGGGGCAT 540
Qy 1403 CACCTGAGAAAGAGCTGATTTTGTATTTTCAGTTTGAAGAAATACTGAAATTTT 1462
Db 541 CACCTGAGAAAGAGCTGATTTTGTATTTTCAGTTTGAAGAAATACTGAAATTTT 600
Qy 1463 CCTTCCTCTGAGATCCCAATTTGTTACATTTTCTCAGGATCTCTAATTT 1642
Db 721 CCTTCCTCTGAGATCCCAATTTGTTACATTTTCTCAGGATCTCTAATTT 780
Qy 1643 CTCTCAATTTCAATTTGTTTAAATTTTACCTCTGGGCTCAATAAGGGCATCTGTGAGAA 1702
Db 781 CTCTCAATTTCAATTTGTTTAAATTTTACCTCTGGGCTCAATAAGGGCATCTGTGAGAA 840
Qy 1703 ATTGGAAGCATTTAGAAATCTTTTGATTTTCCGTGGTTTAAAGGCAATTAAGATG 1762
Db 841 ATTGGAAGCATTTAGAAATCTTTTGATTTTCCGTGGTTTAAAGGCAATTAAGATG 900

```
OY 1763 AGCTTATTACTGGGGTGAGGACAGCTTACTCCATTGTGACCAGATTGTTGGCTAACACA 1822
    |||||
Db 901 AGCTTATTACTGGGGTGAGGACAGCTTACTCCATTGTGACCAGATTGTTGGCTAACACA 960
    |||||
OY 1823 TCCCGAAGATGATTTTGTCAAGAAATTATGTTAATTAATAATATTCAGGATATTTT 1882
    |||||
Db 961 TCCCGAAGATGATTTTGTCAAGAAATTATGTTAATTAATAATATTCAGGATATTTT 1020
    |||||
OY 1883 CCTCTACATAAAGTACCAAT 1903
    |||||
Db 1021 CCTCTACATAAAGTACCAAT 1041
    |||||
```

Search completed: August 23, 2004, 15:19:09
Job time : 7428 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 23, 2004, 19:21:06 ; Search time 471 Seconds
(without alignments)
2854.425 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208
Sequence: 1 MSDHGDVSLPPEDRVRLASQ.....FCSCSHVTVDRAVTTIDLR 424

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
Delop	6.0 , Delext 7.0

Searched: 5669501 seqs, 1585417693 residues

Total number of hits satisfying chosen parameters: 11339002

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xih
-Q=/cgn2_1/USPTO.spool/US09831452/rnuc.17082004_151202_23722/app.query.fasta_1.583
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=p2n.rnpn -MIMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=445 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=us09831452 @cgn 1 1 177 @rnuc.17082004_151202_23722
-NCPU=6 -ICPU=3 -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOC=100
-LONGLOC -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELEXT=7

Database :

Pending Patents NA_New:*
1: /cgn2_6/prodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/prodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/prodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/prodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/prodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq2:*
8: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq3:*
9: /cgn2_6/prodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	2208	100.0	2045	7	US-10-021-698A-3931
2	2208	100.0	2107	7	US-10-784-004-596
3	2208	100.0	2107	7	US-10-784-004-1024
4	1913	86.6	1544	7	US-10-784-004-224
5	1913	86.6	1544	7	US-10-784-004-875
6	1229.5	55.7	4052	7	US-10-370-715B-115
7	810	36.7	1123	7	US-10-021-698A-3934
8	758	34.3	1041	7	US-10-860-790-18
9	758	34.3	1043	7	US-10-860-790-19

10	529	24.0	1972	7	US-10-767-701-13735	Sequence 13735, A
11	521.5	23.6	2559	6	US-10-425-115-126539	Sequence 126539, A
12	498	22.6	1660	6	US-10-425-115-13930	Sequence 13930, A
13	465.5	21.1	8639	5	US-09-404-520B-5634	Sequence 5634, Ap
14	445	20.6	730	7	US-10-779-543-3954	Sequence 3954, Ap
15	411.5	18.6	909	7	US-10-767-701-8700	Sequence 8700, Ap
16	395.5	17.9	539	5	US-09-404-520B-22974	Sequence 22974, A
17	388	17.6	701	6	US-10-425-115-126538	Sequence 126538, A
18	387	17.5	277	7	US-10-793-479-2838	Sequence 2838, Ap
19	350.5	15.9	1674	6	US-10-425-115-164175	Sequence 164175, A
20	335.5	15.2	3709	6	US-10-425-115-93781	Sequence 93781, A
21	321.5	14.6	640	6	US-10-425-115-126535	Sequence 126535, A
22	294	13.3	386	7	US-10-779-543-15310	Sequence 15310, A
23	286	13.0	1335	1	PCT-US03-41761-42714	Sequence 42714, A
24	286	13.0	1335	1	PCT-US03-41761-42714	Sequence 42714, A
25	266	13.0	1335	1	PCT-US03-41766A-42714	Sequence 42714, A
26	275	12.5	505	7	US-10-021-698A-3932	Sequence 3932, Ap
27	257.5	11.7	387	7	US-10-779-543-12862	Sequence 12862, A
28	256	11.6	84504	9	US-60-582-609-20014	Sequence 20014, A
29	252	11.4	436	7	US-10-021-698A-3936	Sequence 3936, Ap
30	225.5	10.2	416	7	US-10-779-543-14802	Sequence 14802, A
31	221.5	10.0	1090	7	US-10-781-469-125	Sequence 125, App
32	210	9.5	557	7	US-10-021-698A-3938	Sequence 3938, Ap
33	204	9.2	201	9	US-60-582-609-90937	Sequence 90937, A
34	203	9.2	1022	7	US-10-767-701-10361	Sequence 10361, A
35	188.5	8.5	397	7	US-10-779-543-14822	Sequence 14822, A
36	181	8.2	201	9	US-60-582-609-90954	Sequence 90954, A
37	159	7.2	117	7	US-10-708-204-6499	Sequence 6499, Ap
38	135.5	6.1	1531974	6	US-10-915-727-12207	Sequence 12207, A
39	134.5	6.1	865	7	US-10-767-701-8859	Sequence 8859, Ap
40	133	6.0	7390	9	US-60-582-609-101	Sequence 101, App
41	131.5	6.0	4266	7	US-10-021-698A-2542	Sequence 2542, Ap
42	131.5	6.0	7151	6	US-10-820-583A-15	Sequence 15, App
43	130.5	5.9	7433	9	US-60-582-609-100	Sequence 100, App
44	130	5.9	10174	9	US-60-592-408-104	Sequence 104, App
45	129.5	5.9	7495	6	US-10-820-583A-16	Sequence 16, App

ALIGNMENTS

RESULT 1
US-10-021-698A-3931/C
Sequence 3931, Application US/10021698A
GENERAL INFORMATION:
APPLICANT: KETIH, TIM
APPLICANT: LITTLE, RANDALL
APPLICANT: VAN EERDEWEGH, PAUL
APPLICANT: DUPUIS, JOSEF
APPLICANT: DEL MASTRO, RICHARD
APPLICANT: SIMON, JASON
APPLICANT: ALLEN, KRISTINA
APPLICANT: PANDIT, SUNIL
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 2976-404US1
CURRENT APPLICATION NUMBER: US/10/021,698A
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/211,749
NUMBER OF SEQ ID NOS: 6160
SOFTWARE: PatentIn 2.1
SEQ ID NO 3931
LENGTH: 2045
TYPE: DNA
ORGANISM: Homo sapiens
US-10-021-698A-3931
Alignment Scores:
Pred. No.: 1.89e-231
Score: 2208.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 2045
Matches: 424
Conservative: 0
Mismatch: 0
Indels: 0

DB: 7 Gaps: 0

US-09-831-452-1 (1-424) x US-10-021-698A-3931 (1-2045)

QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20

DB 1929 ATGTCTGACCATGAGATGAGCTCCGCGCCGAGACCGGATGAGGCTCTCTCCAG 1870

QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40

DB 1869 CTGGGAGATGCGGTAGAGGTGAAGATGAAGACATCCACCCGCTGGACTCCGCTCGGA 1810

QY 41 ValGluIleIleLeuArgMetAlaSerIleTyrSerGlnGluGlyAsnIleGluHisAlaPhe 60

DB 1809 GTTGAGATTATCCGAATGGCATCTTACTCTGAGAGAGGCAACATTAACATGCTTC 1750

QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80

DB 1749 ATCTCTATATAAGATATATCAAGCTTTATTGAGAACTACCAAAACATCGAGATTAC 1690

QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100

DB 1689 AAATCTGCTCTCATCTCTGAAAAAGAAACACAGTAAGAAATTAAGAGATTGATT 1630

QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGln 120

DB 1629 CCCAAAGCAAGAGCTGAAAGCGAGACTTTAAAGATATCCAAAGATATACAGAA 1570

QY 121 TyrAsnGluGluLysLysGluValGluGluLeuAlaArgAsnMetAlaIleGlnGln 140

DB 1569 TATATGAGAAAAAGAAAGAGAGAGAGAAATGGCCCGGAACTGGCCATCGACGCA 1510

QY 141 GluLeuGluLysGluLysGluArgValAlaGlnGluLysGlnGlnGlnGlnGlnGln 160

DB 1509 GAGCTGAAAAAGAAAAACAGAGGTAGCAACACAGAGAGCAATTTGAAACAGAA 1450

QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuLysGluArgLeuLys 180

DB 1449 CAGTTTCATGCTTCGAGAGATGATCGGAAACAGAGCTTAAGAAAAAGCACTGAAA 1390

QY 181 IleValGlnGluPheGlyLysValAspProGluLeuGlyLysProLeuValProAspLeu 200

DB 1389 ATTGTACAGAGTTTGGAAAGTAGACCTGGCTGAGTGGCCGCTGAGCTGACTTG 1330

QY 201 GluLysProSerLeuAspValPheProThrIleThrValSerSerIleGlnProSerAsp 220

DB 1329 GAGAAAGCTCTCTAGATGTGTTCCCACTTAACATCTCATCCATACAGCTTCAGAC 1270

QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240

DB 1269 TGTCAACAACACTGTAAGGCGCAGCTAACCACTGTGTGACAGGTGCTTGAAACCTGGA 1210

QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro 260

DB 1209 GCACTGAGCAACTAGAAAGATTTCCCAATCATGATGATGGCCCATGTGGTGCTT 1150

QY 261 GlyArgLeuLysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280

DB 1149 GGGCGGCTGTCCCAAGTTTCTCCAGTTAGCCAGTCCAAACTGCCCCGGAGAGTGAG 1090

QY 281 ThrCysGlyIleLeuLysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300

DB 1089 ACATGTGAATTTCTGTGAAAACTGATGAGAAATTAATTTACCATTTACCATTTCTTC 1030

QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320

DB 1029 ATCCCAAGCAAGAGCTGGGTCTGATTACTGCAACACAGAGAGCAAGAAAGAACTTTC 970

QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrPheIleThrHisIleProThrGln 340

DB 969 CTCAATACAGGATCGACAGGCTCATCACTGGAGCTGATCTACTCATCCCAACAG 910

QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360

DB 909 ACCGCGTTCTCTCCATGTGACCTTACACATCACTGCTCTTACAGAGATGATGTCGA 850

QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380

DB 849 GAGTCAGTACCATTTCTTCTCTCCCAAGTTCCAGAAATCGATTTCTTAACACT 790

QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400

DB 789 GACCATGACCTAGAGAGATTTCTTCTCTGCGCAGAAAGATTTTATCCACAGCAAG 730

QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420

DB 729 GATCCACCTCTGTCTGTAGCTGACGACGATGATGTGTGACAGAGAGTACCATC 670

QY 421 ThrAspLeuArg 424

DB 669 ACAAGCTTCA 658

RESULT 2

US-10-784-004-596

/ Sequence 596, Application US/10784004

/ GENERAL INFORMATION:

/ APPLICANT: Biogen Idec

/ TITLE OF INVENTION: Surrogate Markers of Pain

/ FILE REFERENCE: 08201.6029-00000

/ CURRENT APPLICATION NUMBER: US/10/784,004

/ CURRENT FILING DATE: 2004-02-20

/ NUMBER OF SEQ ID NOS: 1251

/ SOFTWARE: PatentIn version 3.2

/ SEQ ID NO 596

/ LENGTH: 2107

/ TYPE: DNA

/ ORGANISM: human

US-10-784-004-596

Alignment Scores:

Pred. No.: 1,996-231 Length: 2107

Score: 2208.00 Matches: 424

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 7 Gaps: 0

US-09-831-452-1 (1-424) x US-10-784-004-596 (1-2107)

QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20

DB 188 ATGTCTGACCATGAGATGAGCTCCGCGCCGAGACCGGATGAGGCTCTCTCCAG 247

QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40

DB 248 CTGGGTAGTGGCGGTAGAGGTGAATGACATTTCCACCCGCTGGTACTTCGCTCGA 307

QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGlnGluGlyAsnIleGluHisAlaPhe 60

DB 308 GTTGAGATTATCCGAATGCAATCCATTTACTCTGAGAGAGCAACATTAACATGCCCTC 367

QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80

DB 368 ATCTCTATATAAGATATATCAAGCTTTATTGAGAACTTACCAAAACATCGAGATTAC 427

QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuLysGluIleAlaPhe 100

DB 428 AAATCTGCTCTCATCTCTGAAAAAGAAACACAGTAAGAAATTAAGAGATTGCAATT 487

QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGln 120

DB 488 CCCAAAGCAAGAGACTGAGGCGAGAGCTGTTTAAACATATATCCAAAGAAATATACAGAA 547

QY 121 TyrAsnGluGluLysLysGluValGluGluLeuAlaArgAsnMetAlaIleGlnGln 140

DB 548 TATATGAGAAAAAGAAAGAGAGAGCAAGAAATTTGCGCCGGAACATGCGCATCCAGCA 607

Qy	141	GlulLeuGlnLysGluLysGlnArgValAlaGlnGlnLysGlnGlnGlnGlnGlnGln	160
Db	608	GAGCTGAGAAAAAGAAAAAGAGAGGTAGCAACAAGAAAGACAGACATTTGAAACAGGAA	667
Qy	161	GlnPheHisAlaPheGluGluMetLeuArgAsnGlnGluGluGlnLysGluArgLeuLys	180
Db	668	CATTTCATGCTCTTCAGAGAGATGATCCGAAACAGAGCTTAGAAAAAGAGCACTGAAA	727
Qy	181	LeuValGlnGluPheGluLysValAspProGluLysGluLysProLeuValProAspLeu	200
Db	728	ATTGTACAGAGATTGGAGAGGTAGACCTCGCTGAGTGGCCGCTAGTGCTGACTTG	787
Qy	201	GluLysProSerLeuAspValPheProThrLeuThrValSerSerLeuGlnProSerAsp	220
Db	788	GGAAGCCCTCCCTAGATGTGCTCCCACTTAACGTCTCAATCAATACAGCTTATGAC	847
Qy	221	CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly	240
Db	848	TGTCAACAACATGTAGAGCGCAGCTAGCAACCTGTGTGACAGAGCTTGAACCTGGA	907
Qy	241	AlaLeuSerAsnSerGluSerLeuProThrLeuAspGluLeuArgHisValValPro	260
Db	908	GCACTAGACAACTCAGAAAGTATTTCCCAATGTAGATGGCATGTGGTGGTGGCT	967
Qy	261	GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu	280
Db	968	GGGCGGCTGTGGCCACAGTTCTCCAGTTAGCCAGAGCCAAACCTCGCGGAGTGGAG	102
Qy	281	ThrCysGlyTLeuLysGlyLysLeuMetArgAsnGlnPheThrTLeuThrHisValLeu	300
Db	1028	AATGTGTGAATTTCTCTGTGAATACTGATGAGAAATGAATTTACATTAACCATGTTTC	1087
Qy	301	LeuProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLysLeuPhe	320
Db	1088	ATCCCCAAGCAAAAGTGCTGGTGTGTACTGTACTGCAACAGAGAAAGAAACATTTTC	1147
Qy	321	LeuIleGlnAsnGlnGlnGlnLysLeuIleThrLeuGlyTProIleHisThrHisProThrGln	340
Db	1148	CTCATACAGATCAGCAGGCGCTCATCAACATGGGTGATTTCAATCAACCCCAACAG	1207
Qy	341	ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro	360
Db	1208	ACCGGTTTCTCTCCAGTGTGCACCTACCACTCACTGCTCTTACCAAGATGTTGGCA	1267
Qy	361	GluSerValAlaIleValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuThr	380
Db	1268	GAGTCAGTAGCCATGTTGTGCTCCCCCAAGTCCAGGAAACGAGATTTCTTAAACTACT	1327
Qy	381	AspHisGlyLeuGlnGlnLysSerSerCysArgGlnLysGlyPheHisProHisSerLys	400
Db	1328	GACCATGAGCATGAGAGAGATTCTTCCTGCTGGCCAGAAAGAGATTTCAATCCACACGCAAG	1387
Qy	401	AspProPheLeuPheCysSerCysSerHisValThrValValAspArgAlaValThrIle	420
Db	1388	GATTCACCTCTGTTTGTAGTGTGACCAACCTGACTGTTGTGGACAGAGCAGTGCACATC	1447
Qy	421	ThrAspLeuArg	424
Db	1448	ACAGACCTTCGA	1459
RESULT 3			
US-10-784-004-1024			
Sequence 1024, Application US/10784004			
GENERAL INFORMATION:			
APPLICANT: Biogen Idec			
TITLE OF INVENTION: Surrogate Markers of Pain			
FILE REFERENCE: 08201.6029-00000			
CURRENT APPLICATION NUMBER: US/10/784,004			
CURRENT FILING DATE: 2004-02-20			
NUMBER OF SEQ ID NOS: 1251			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO 1024			

:	LENGTH:	2107
:	TYPE:	DNA
:	ORGANISM:	human
:	US-10-784-004-1024	
 Alignment Scores:		
Pred. No.:	1.99e-231	Length: 2107
Score:	2208.00	Matches: 424
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	7	Gaps: 0
 US-09-831-452-1 (1-424) x US-10-784-004-1024 (1-2107)		
QY	1 MetSerSPHISGIYAPVALSERLEUPROPGIUAASPAAGVALAALAEUSERGIN	20
Db	188 ATGCTACCAATGAGATGTGAGCCCTCCGCCGAAGACCGGGTGAGGCCTCTCCAG	247
QY	21 LeugIYSERIALAYAGUVALANSIUASPIIAEPORCARXATGYPhaIXSergly	40
Db	248 CTGGTGATCGGTGAGTGAAGTGAAGACATTCACCCCGGGTACTCTCGCTTCGGA	307
QY	41 ValInUIleIlleaRgMetAlaserIIeYrSergIngUGlUYasnlIEglUHIsalApe	60
Db	308 GTTAGAGATTTCGAAATGGCATCATTACTCTGAGAAGGACAATTGAACATCGCTTC	367
QY	61 ILeuutYrsnlySTyrlIEtRIleuPheIIegUlYSLeUProLYSHIsrgAsPTyr	80
Db	368 ATCTCTATAACAAGATATATCACGCTCTTTATVGAAAATCAACMAACATCGAGATTAC	427
QY	81 LysSErAlaValIleEProGluYSLysASPThraVILysLSLeuYLGLIleAlaPhe	100
Db	428 AAATCTGCTCTACTCTCTGAAAAAGAACACAGTAAGAATTAAAGAGATTGCTATT	487
QY	101 ProLYsaIaeGUleuleuLYsaIaeGUleuleuLYsaTYrThryLsgUTYrThrgLU	120
Db	488 CCCMAACCAAMACGCTGGAAGCAGCGCTGTAAAAACATATACCAAAGAATATACGAA	547
QY	121 TyrAsmGUglULysLYSGluIaeGUleuAlaAgasmeAlaIleGln	140
Db	548 TATATCAAAAAGAAAGAAAGAACAGACAGAAATTGCCCGGAACATGCCCATTCCAGAA	607
QY	141 GlueUGluYSGluYSLmaRYVaIAagInGluYSGngInGLEUGluGlnLU	160
Db	608 GAGCTGGAAGAAAGAAACAGAGGTTAGACACACAGACGACGACATTGGAACAGAA	667
QY	161 GlnPHHisalApeGUleuMeTleArghsngInGUleuGluYSGluAgleuLYS	180
Db	668 CAGTTCCATGCTTCAGAGAGATGATCCGGAAACCGAGGCTAGAAAAAGCGACTGAAA	727
QY	181 IleValGlnGUlPheGlyLYsValASPProGlyLeuGlyGPoleuValProASPLeu	200
Db	728 ATTGTACAGAGATTGGAAAGTTAACCTTGCCCTAGGTGGCCCGCTAGTCTGACTTG	787
QY	201 GlULYPrOSerILEuaSPvaIPheProThriLeuthraValseSerllEglnProserASP	220
Db	788 GAAGAACCCCTCTTGAAGATGTCCCAACCTTAACAGCTTCATCAATACAGCTTCAGAC	847
QY	221 CySHISThThraLPProAlaYSPProProvalValaSParSereuLYSPGOJY	240
Db	848 TGTCACACAACGTAAAGCCACGCTAGCCACCTGTGTGACAGGTCTTGAAACCTGGA	907
QY	241 AlaLeuSeArSseGluSeRIleProThrilEaspGlyLeuAryhisValaValaPro	260
Db	908 GCACCTAGCACTCGAAGAGATTCCCAATCATCATGATTGGCCACTGGGTGGTGCCT	967
QY	261 GLYAArgLeuCYSPGlnPHeuGluIleuAlaserlaasnthraLaargGlyValIGlu	280
Db	968 GGCGGCTGTGCCCAAGTTTCTCOAGTTAGCCAGTGTCCAACACTGCGGGGAGTGGAG	1022
QY	281 ThrCyGlyIleLeuCYGlyLYleuWeCaArgsmGUlPheThrilEthrHisValleu	300

Db 1028 ACATGTGATCTCTGTGAAACGTAGAGATGATTTACCATTCATGTTCTC 1087
 QY 301 ILEPRLYSGLINSEKALIGLYSERAPRYRQYSAANTHRIUANGIUGIULPHE 320
 Db 1088 ATCCCAAGCAAGAGTGTGTGATTAATTCACACAGAAAGAAAGAACTTTC 1147
 QY 321 LEUILEGINASPGLINGIULYLEUILETHREUGIYTRILEHISTHISPROTHRI 340
 Db 1148 CTCATACAGATCAGAGGGCTCATCACTGGGCTGATTCACTCACCCACAG 1207
 QY 341 THRAlAPHEUSERSERVALASPHEHISTHISCYSEERYRGINMETHELEUPRO 360
 Db 1208 ACCGCTTCTCTCCAGTGTGACCTACACTCATCTTACAGATGATGTTCCA 1267
 QY 361 GLUSERVALALALEVALCYSEERPROLYSPHEGNIURHGIYPHEPELYSELEUTH 380
 Db 1268 GAGTCAGTGCATGTTGCTCCCCCAAGTCCAGAACTGATTTTAAACTACT 1327
 QY 381 ASPHISGLYLEUGIUGIULYLEUSERCYSAARGIULYSGIYPHEHISPROHISSELYS 400
 Db 1328 GACCATGACTAGAGAGATTTCTTCCTGCGCCAGAAAGATTTCAATCCACAGCAG 1387
 QY 401 ASPPROLEUPHECYSEERCYSEERHISVALTHRYVALASPARGALAVALTHTLE 420
 Db 1388 GATCCACCTCTGTCTGTAGTGCAGCCAGCTGACTGTTGACAGAGCAGTACCATC 1447
 QY 421 THRASPHEUARG 424
 Db 1448 ACAGACCTTCGA 1459

RESULT 4
 US-10-784-004-224
 ; Sequence 224, Application US/10784004
 ; GENERAL INFORMATION:

; APPLICANT: Biogen Idec
 ; TITLE OF INVENTION: Surrogate Markers of Pain
 ; FILE REFERENCE: 08201.6025-00000
 ; CURRENT APPLICATION NUMBER: US/10/784.004
 ; CURRENT FILING DATE: 2004-02-20
 ; NUMBER OF SEQ ID NOS: 1251
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 224
 ; LENGTH: 1544
 ; TYPE: DNA
 ; ORGANISM: rat
 ; US-10-784-004-224

Alignment Scores:

Pred. No.: 3e-199 Length: 1544
 Score: 1913.00 Matches: 357
 Percent Similarity: 91.98% Conservative: 33
 Best Local Similarity: 84.20% Mismatches: 34
 Query Match: 86.64% Indels: 0
 Gaps: 0

US-09-831-452-1 (1-424) X US-10-784-004-224 (1-1544)

QY 1 MetSerAPHISGLYASPVASERLEUPROGLUASPARGVALARGALAEUSERGIN 20
 Db 51 ATGTCTGACCATGCGAGTGTGAGCTCCACCCCAAGACCGGGTGGAGATTTGTGCGAA 110
 QY 21 LeuGISERRALAVAGIULVALASNGIULASPILEPROFARGARGIYRPHARGSERGIY 40
 Db 111 CTCGGTAGTGCAGTGTGAGTTAAATGAAAGACATTCGCCCGCTGCTCTTCGGGT 170
 QY 41 ValGIULIIELEARGMETALASERILEYSEERGIUGIULYASNIIEGIUHSALAPE 60
 Db 171 GTTGAGATCATCCGATGCGATTCATTTACTCTGAAAGAGCAACATTGAAACATGCTTT 230
 QY 61 ILEUITYRANLYEYRIETHTLEUPHEHIEGIULYSEUPROLYSHISARGAPRYR 80
 Db 231 ATCTCTACAAACAGTACATCAGCTGTATTGTAATAAACTCCAAAACCCGAGACTAC 290

QY 81 LYSSEKALVALIIEPROGLULYSELYSEAPRTHRYVALYSLYSEULYSGIULIIEALPHE 100
 Db 291 AAATGGCCATCATTTCCCGAAGAAAGACCGCGTCAAGAAATTTAAAGATGTCCTTTC 350
 QY 101 PROLYSALAGIUGIULYSELYSALIGIULYUENLYSEYRGTHTHYSGIUTRYTHRI 120
 Db 351 CTTAAAGGGAGAGCTGAAGACAGAACTCTTGAAGATACCCAAAGAGTGTGAGCAG 410
 QY 121 TYRANGIUGIULYSELYSGLUALAGIUGIULYUENLARGASMECALALEGININ 140
 Db 411 TATAAGGCCAAGAAAGAAAGAAAGAGAGAACTTCCGAAATATCCGATCCAGCAA 470
 QY 141 GIULYUGIULYSGIULYSELYSVALALAGIUGIULYSGIULYSGIULYSGIULY 160
 Db 471 GAACGGGAAAAAGAAAGCAAGAGTGTGACAGAGAAAGAGAGAGAGAGCTCGAGCAG 530
 QY 161 GINPHEHISALPHEGIUGIUMETILEARGASNGIULYUENGIULYSGIULYSELYS 180
 Db 531 CAGTCCATGCTCTTGAGAAATGATATCAGACAGAGAGCTTAAAGAAAGCGGCTAAA 590
 QY 181 ILEVALINGIULPHEGIULYSELYSVALASPARGIYUENGIYGLIYPROLEUVALPROAPLEU 200
 Db 591 ATGTTCAGAGTTCGGGAGAGTGAACCTGCGCCGCGCGGCTGCTGCTGCTGATCTG 650
 QY 201 GIULYSPROSERLEUASPVALPHEPROTHRYVALSERIIEGINPROSEKAP 220
 Db 651 GAAAAGCCCTGTATGATGTGCGCCCATTCACCTTTCGCCCAAGAGATTCAGAC 710
 QY 221 CYSHISTHRYTHRYVALARGPROALALYSPROVALVALASPARGSERLEULYSPROGLY 240
 Db 711 TGTAAACAACCCCGAGGCGAGCTAAGCCACTGTGTGACAGTCCCTCAAACTCGGA 770
 QY 241 ALAEUSERAUSERGUSERILEPROTHRIEASGLIYUENYRHSIVALVALPRO 260
 Db 771 GCATTAAAGCGATAGAAATGTTCCACATTGAAGCTGCGCCACATGTGTGCGCC 830
 QY 261 GLYARGIUCYSPROGINPHEUGIULYUENLASERIALASNTHRYALARGIYVALGIU 280
 Db 831 CGCAATCTGTGCTCAGATTTCTCCAGCTTCCAGCGCCCAACCTCCCAAGGACATCGAG 890
 QY 281 THRCYSGIYILEUUCYSGIULYSELYSEUMERFARGANGIUPHETHRIETHRHSIVALLEU 300
 Db 891 ACCTGTGAGTCTCTGTGAAAACCTGATGAGAAATGAAATTCACAACTCACATGTTCTC 950
 QY 301 ILEPROLYSGINSEKALIGLYSERAPRYRQYSAANTHRIUANGIUGIULPHE 320
 Db 951 ATCCCAAGCAAAATGTGTGGCTGATTTATTCACACAGAGATGAAAGAAATTTTC 1010
 QY 321 LEUILEGINASPGLINGIULYLEUILETHREUGIYTRILEHISTHISPROTHRI 340
 Db 1011 TTTATGACAGATGATCTTGAGATCTTCACCTTGGCTGATGACACCCATCCACCCAA 1070
 QY 341 THRAlAPHEUSERSERVALASPHEHISTHISCYSEERYRGINMETHELEUPRO 360
 Db 1071 ACCGCTTCTCTCCAGTGTGAGTGTGACACACACACCTGCTCTACCAAAATGATGTTACCA 1130
 QY 361 GLUSERVALALALEVALCYSEERPROLYSPHEGNIURHGIYPHEPELYSELEUTH 380
 Db 1131 GAGTTCATGACATATGCTCTCTCCCAAGTTCCAGAGACATGATCTTTAAATTAAT 1190
 QY 381 ASPHISGLYLEUGIUGIULYLEUSERCYSAARGIULYSGIYPHEHISPROHISSELYS 400
 Db 1191 GACTATGCTCTCAAGAGATTTCAACCTGCGGCAAGAAAGGCTTTTCAACCCCATGACAGA 1250
 QY 401 ASPPROLEUPHECYSEERCYSEERHISVALTHRYVALASPARGALAVALTHTLE 420
 Db 1251 GACCAACCGCTGTCTGTGACTGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1310
 QY 421 THRASPHEUARG 424
 Db 1311 ACAGACCTTCGA 1322

RESULT 5

```
US-10-784-004-875
; Sequence 875, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6028-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 875
; LENGTH: 1544
; TYPE: DNA
; ORGANISM: Rat
US-10-784-004-875

Alignment Scores:
Pred. No.: 3e-199      Length: 1544
Score: 1913.00      Matches: 357
Percent Similarity: 91.98%      Conservative: 33
Best Local Similarity: 84.20%      Mismatches: 34
Query Match: 86.64%      Indels: 0
DB: 7      Gaps: 0

US-09-831-452-1 (1-424) x US-10-784-004-875 (1-1544)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspAlaArgAlaLeuSerGln 20
Db 51 AAGTCGACATGCAGATGAGCGCTCCACCCAGACCGGGAGGATTCCTCCCAA 110
QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgGlyArgPheArgSerGly 40
Db 111 CTCGGTAGTGAGTGAATTAAATGAACACTTCGCCGCCCTCCCTACTTTCGTCGGGT 170
QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluIleGluAlaPhe 60
Db 171 GTTAGATCATCCCGCATGGCATTCATTACTGAGAGGACCAATTGAACATGCTTT 230
QY 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluIleLeuProIleHisArgAspTyr 80
Db 231 ATCCCTTACAAACAAGTACATACCGCTGTTATTAATAAACTTCCAAACACCGGACTAC 290
QY 81 LysSerAlaValIleProGluIleLysAspThrValIleLysLeuLysGluIleAlaPhe 100
Db 291 AAATCGGCATCATTCGCCGAGAGAAAGACGCGTCAAGAAATTAAGANTGTCGCTTC 350
QY 101 ProLysAlaGluIleLysValGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
Db 351 CCTAAACGGAGAGCTGAAGACAGAACTCTTGAAGAGATACCAAAAGATGAGCAG 410
QY 121 TyrAsnGluIleLysLysGluAlaGluIleLeuAlaArgAsnMetAlaIleGlnGln 140
Db 411 TATAAGAGCGAAAGAGAGAGAGAGAGAACTTCGCCGAATATTCGCCATCCAGAA 470
QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnIleGlnGlnGlu 160
Db 471 GAACCTGAAAAAGAAAAAGACAGAGAGTGCACAGACAGAGAGAGAGAGCTGAGCAGAG 530
QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnIleGluGluLysGluArgLeuLys 180
Db 531 CAGTTCATGCTCTTGAGAGAGATGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590
QY 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyLysProLeuValProAspLeu 200
Db 591 ATTGTTCAAGAGTTCGGAGAGTAGACCTCGCGCGCGGCTCGCTCCCTGATCTG 650
QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
Db 651 GAAGAGCCCTGCTGAGATGAGCCGCCCAAGTTCACTTTCTCGGCCAGCAGACTTCAGAC 710
QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
Db 711 TGTAACACAAACCTCGAGCGCAGCTAAGCCACTGTGTGAGACAGTTCCTCAAACTGGA 770
```

```
QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAsnGlyLeuArgHisValAlaValPro 260
Db 771 GCATTAAAGCTCATAGAAATGTTCCACCATTTGAGAGCGCTGCGCCACATTTGTGTGCC 830
QY 261 GlyArgLeuGlySerProGlnPheLeuGlnIleValAsnThrAlaArgGlyValGlu 280
Db 831 CGCAATCTGTGTGAGATTTCTCCAGCTTGCAGCGCCAGCCACACTGCCAAAGCATGAG 890
QY 281 ThrCysGlyIleLeuGlySerGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
Db 891 ACCGTGAGTCTCTGTGGAAAACTGATGAGAAATGAAATTCACAACTCACACATGTTCTC 950
QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
Db 951 ATCCCGACAGAAATGGTGGCCCTGATTTATGGCACACAGAAATGAAAGAAATTTTC 1010
QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
Db 1011 TTATATGAGATGATCTTGGACATCTCCACTCTTGCTGATCCACACCATCCAAACCAA 1070
QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPro 360
Db 1071 ACGGCTTTCTGTCCAGTGTGATCTGACACACGACGCTCTCTACCAATGATGTACCA 1130
QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380
Db 1131 GAGTCCATGACATTTCTGTCTGCTCCCAAGTTCCAGAGAGCTGATTTTAATTAACT 1190
QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
Db 1191 GACTATGCGCTTCAAGAAATTCACACTGCGGAGAAAGGCTTTACCCCACTGACAGA 1250
QY 401 AspProProLeuPheCysSerCysSerHisValThrValValAspArgAlaValThrIle 420
Db 1251 GACCCACCGCTGTCTGTGACTGACGACCATGTCACTGTAAGACAGAAATTTGACGATC 1310
QY 421 ThrAspLeuArg 424
Db 1311 ACAGACCTTGA 1322

RESULT 6
US-10-370-715B-115
; Sequence 115, Application US/10370715B
; GENERAL INFORMATION:
; PatIn Docket: Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; FILE REFERENCE: P194891-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 115
; LENGTH: 4052
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-370-715B-115

Alignment Scores:
Pred. No.: 1.63e-123      Length: 4052
Score: 1229.50      Matches: 232
Percent Similarity: 73.97%      Conservative: 72
Best Local Similarity: 56.45%      Mismatches: 92
Query Match: 55.68%      Indels: 15
DB: 7      Gaps: 2
```

```

US-09-831-452-1 (1-424) X US-10-370-715B-115 (1-4052)
QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
Db 872 ATGCTGACCATACATACATGTTTCCCTTAAGCCAGAAAGAGGAGTCCGTCCTTAAGCAAG 931
QY 21 LeuGlySerAlaValGluValaGluAspGluLeuProProArgArgGlyThrLeuSerGly 40
Db 932 CTGGTGTGATATCACCACATGAGTGAAGACATCACTCCACGAGGTTACTTAAAGTGTGA 991
QY 41 ValGluLeuLeuArgMetAlaSerIleIleIleGluGluGlyAsnIleGluHisAlaPhe 60
Db 992 GTTAGAGATGAGAGAGATGAGCTGCTGTGTATTTGGAAGAGAAATTTGGAATATGCTTT 1051
QY 61 IleuLeuTyrosLeuIleuLeuLeuPheIleGluLeuLeuProIleuLeuPheIleuLeu 80
Db 1052 GTTCTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 1111
QY 81 LysSerLeuAlaValIleuProGluLeuLeuAspThrValLysLeuLeuGluLeuAlaPhe 100
Db 1112 CAGCAATGTGACGACTGTAAGAGAGAGAGATATATATATATATATATATATATATATAT 1171
QY 101 ProLysAlaGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
Db 1172 CCAAGGACAGATGATATGAAAAACGACCTTTAAAGAAATATATATATATATATATATATAT 1231
QY 121 TyrAsnGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140
Db 1232 TATTTGCAAGCAAAAAACAATATATATATATATATATATATATATATATATATATATAT 1291
QY 141 GluLeuGluLeuGluLeuGluLeuGluLeuGluLeuGluLeuGluLeuGluLeuGluLeu 160
Db 1292 TTGATATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1351
QY 161 GluPheHisAlaPheGluGluMetIleArgAsnGluGluLeuGluLeuGluLeuGluLeu 180
Db 1352 CAGTTTCTGTTTTCGAAGATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411
QY 181 IleValGluGluLeuPheGlyLeuValAspProGlyLeuGlyLeuProLeuValProAsp 200
Db 1412 AGTCAGCAAACTCAAGG-----CTGTGAGAGAGAGAT 1444
QY 201 GluLysProSerLeuAspValPheProThr-----LeuThrValSer 214
Db 1445 GATGGAGAGCGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1504
QY 215 SerIleGlnProSerAspCysHisThrThrValArgProAlaLysProProValAlaAsp 234
Db 1505 GCACATCAACCTATATAAAAGTATGCAACCAATTATGCTTACCTCTCTCTCTCTCTCT 1564
QY 235 ArgSerLeuLysProGlyAlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeu 254
Db 1565 AGGAGCTTAAGCGCAGCTGCTAAGCTAAGCTGCTTCAAGATTAGTGTTGAAGGACTG 1624
QY 255 ArgHisValAlaValProGlyArgLeuCysProGlnPheLeuGlnIleAlaSerAlaAsn 274
Db 1625 CGATGTGATGTTTCCAGAAAGATCTTTGCCAATAATTTGCAACCTGGGCAAGATCTTAAT 1684
QY 275 ThrAlaArgGlyValGluThrCysGlyIleLeuCysGlyLeuLeuMetArgAsnGluPhe 294
Db 1685 ACAGTGAAGAGATGAAGAACTGTGATATCTGTGGAAGAAAGACACATATATGAATTT 1744
QY 295 ThrIleThrHisValLeuIleProLysGlnSerAlaGlySerAspTyrCysAsnThrGlu 314
Db 1745 ACTATTACCCATGATATGCTGCAAGAGAGCTGGGAGCCAGCATATGTCGATGAGAG 1804
QY 315 AsnGluGluGluLeuPheLeuIleGlnAspGlnGlnGluLeuIleThrLeuGlyTyrIle 334
Db 1805 AATGTAGAGAGATTTATCAATGTTCAAGATCAACATGATCTCTCACTCTGATGATGATC 1864
QY 335 HisThrHisProThrGlnThrAlaPheLeuSerSerValaPheHisThrHisCysSer 354

```

```

Db 1865 CATACATCCCACTCAAACTGATTTTATCCAGCGTTGATCTTCACTCACTGTTCC 1924
QY 355 TyrGlnMetMetLeuProGluSerValAlaIleValCysSerProLysPheGlnGluThr 374
Db 1925 TATCACTCATGTTTCCAGAGCGCATTTCTTTGCTTCAACCAAGATTAAGACACT 1984
QY 375 GlyPhePheLeuLeuThrAspHisGlyLeuGluGluIleSerSerCysArgGlnLysGly 394
Db 1985 GGCATCTTCAAGCTCAACCAATGCTGAGATCTTGAGGTTCTGTGTTTAAAAAAGGCG 2044
QY 395 PheHisProHisSerLysAspProProLeuPhe 405
Db 2045 TTTCATCCACACCAAGAGCCAGGCTGTT 2077

RESULT 7
US-10-021-698A-3934/c
Sequence 3934, Application US/10021698A
GENERAL INFORMATION
APPLICANT: KEITH, TIM
APPLICANT: LITTLE, RANDALL
APPLICANT: VAN EERDEWEH, PAUL
APPLICANT: DUPUIS, JOSE
APPLICANT: DEL MASTRO, RICHARD
APPLICANT: SIMON, JASON
APPLICANT: ALLEN, KRISTINA
APPLICANT: PANDIT, SURIL
TITLE OF INVENTION: NOCOTINIDE AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 2976-4044US1
CURRENT FILING DATE: 2001-10-22
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/211,749
NUMBER OF SEQ ID NOS: 6160
SOFTWARE: PatentIn 2.1
SEQ ID NO 3934
LENGTH: 1123
TYPE: DNA
ORGANISM: Homo sapiens
US-10-021-698A-3934

Alignment Scores:
Pred. No.: 2,436-78 Length: 1123
Score: 810.00 Matches: 174
Percent Similarity: 69.49% Conservative: 15
Best Local Similarity: 63.97% Mismatches: 24
Query Match: 36.68% Indels: 59
Gaps: 3

US-09-831-452-1 (1-424) X US-10-021-698A-3934 (1-1123)
QY 16 ArgAlaLeuSerGlnLeuGlySerAlaValaGluValaGluAspIleProProArgArg 35
Db 799 AGAGAGATGGCCAGCTGGGTATTAAGT-----CTCCCTCCGCTGAC 755
QY 36 TyrPheArgSerGlyValGluIleIleArgMetAlaSerIleIleIleGluGluGluGlu 55
Db 754 TTCTGGAACCTTCAATATATAGAGGCTTCAAGATTAAGGAGGTTGATAGGAGTCT 695
QY 56 IleGlu-----HisAlaPheIleuLeuTyrosLeuLeuLeuLeuLeuLeuLeuLeu 71
Db 694 ATGACATTTGACACATTTGCTTCAATATCAATTTGAACCTGTTCTTAAAGCTCTTAT 635
QY 72 GluLysLeuProLysHisArgAspTyrLysSerAlaValaIleProGluLysAspThr 91
Db 634 GAGAACTACCAAAACATGAGATTAACAATGCTGTCATCTCGAAAGAAAGACCCA 575
QY 92 ValLys----- 93
Db 574 GTAAAGTGGGTTTCACTTTCATTTGTTGCCCACATATAAGCTTCAGAGCAGACAGACT 515
QY 93 ----- 93

```

Db 514 GACTTCAGATATAAGTAGTAGTCTGTCATCTGTAGTACAGATCCTGCATATAGAGATA 455
 QY 94 -----Lysleuylsgluilealphepolly 102
 Db 454 AGGCCATTTCACCAACAGTTCTGMAAGTAGCAGAAATTAAGAGATTGCAATTCACCA 395
 QY 102 salagluileuylsalagluileuylsargtyrthrlysgluityrthrlygtyras 122
 Db 394 AGCAGAAAGAGCTGAAAGCAGAGCTGTTAAACGATATACCAAAATATACAGATTAA 335
 QY 122 nglngluylslysgluialagluileualargsnmerlatlelngluile 142
 Db 334 TGAAGAAAGAAAGAAAGAGAGGAAATTTGCCCGGAACAAGCCATCCAGCAAGAGCT 275
 QY 142 ugllyleugluylsargvalalagluileugluileugluileugluileugluile 162
 Db 274 GGAAGAAAGAAAGAGAGGTTAGCAACAGACAGACGCAATTGGAACAGAAACGTT 215
 QY 162 ehialaphegluileucileargsnngluileugluileugluileugluile 182
 Db 214 CCATGCCCTTCAGAGAGATGATCCGGAACAGAGACTGAAAGAGGAGCTGAAATTTGT 155
 QY 182 lglngluiphegllystvalappprogllyleuglyglyproleuvalproaspueugly 202
 Db 154 ACAGAGATTGGGAGAGTAGACCTTGCTAGGGCCGCTAGTGTGCTGACTTGAGAGAA 95
 QY 202 sProSerleuaspvalPheProThrLeuThrValSerSeriIegInProSeraspCysH 222
 Db 94 GCCCTCCTTAGATGTGTTCCTCCACCTTAACATGTTCACTCATACAGCCTTCAGACTGTCA 35
 QY 222 sThThValargProalalysProprovalval 233
 Db 34 CACAACGTGAAGGCCACAGTAAACACACCTGTGTG 1

RESULT 8

US-10-860-790-18
 ; Sequence 18, Application US/10860790
 ; GENERAL INFORMATION:
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Hill, Paul
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.462C11
 ; CURRENT APPLICATION NUMBER: US/10/860,790
 ; CURRENT FILING DATE: 2004-06-02
 ; NUMBER OF SEQ ID NOS: 624
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 18
 ; LENGTH: 1041
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 544
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-860-790-18

Alignment Scores:

Pred. No.: 1,1e-72 Length: 1041
 Score: 758.00 Matches: 140
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 34.33% Indels: 0
 Gaps: 0

US-09-831-452-1 (1-424) x US-10-860-790-18 (1-1041)

QY 285 LeuCySGlyLysLeuMetArgsnGluPheThrIleThrHisValLeuIleProlySgin 304
 Db 1 CTCGTGGAAGAACTGATGAGGAATGAATTATTAACATTACCATGTTCTCATCCCAAGCAA 60

QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPheLeuIleGlnasp 324
 Db 61 AGTGTGGGTCTGATTATTCGCAACACAGAAAGAAAGAACTTTCTCTCATACAGGAT 120
 QY 325 GlnGluGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrSginThrAlaPheLeu 344
 Db 121 CACAGAGGCTCATACATCAGCTGGCTGATTCATACCAACCCACAGACCCGCTTCTC 180
 QY 345 SerSerValaspIleuHisThrHisCysSerTyrGlnMetLeuProGluSerVala 364
 Db 181 TCCAGTGTGACCTACACACTGCTCTTACAGATGATGTGCCAGAGTCAGTAGCC 240
 QY 365 IleValCysSerProlyspheGluGluThrGlyPhePheLeuThrAspHisGlyLeu 384
 Db 241 ATTGTTGCTCCCCCAAGTTCCAGAACTGGATTTTAACTTAACCTAACCACTGAGCTA 300
 QY 385 GluGluIleSerSerCysArgGluysglyPheHisProHisSerIysaspProleu 404
 Db 301 GAGAGATTTCCTTCCTGTCGCGCAAGAAAGATTTCATCCACAGCAAGATCCACTCTG 360
 QY 405 PheCysSerCysSerHisValThrValaIaspAlaGalaValThrIleThrSpleuArg 424
 Db 361 TTCTGTAGCTGAGCCAGCTGATGTTGTGACAGAGCAGTACCATCACAGACCTTCA 420

RESULT 9

US-10-860-790-19
 ; Sequence 19, Application US/10860790
 ; GENERAL INFORMATION:
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Hill, Paul
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.462C11
 ; CURRENT APPLICATION NUMBER: US/10/860,790
 ; CURRENT FILING DATE: 2004-06-02
 ; NUMBER OF SEQ ID NOS: 624
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 19
 ; LENGTH: 1043
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-860-790-19

Alignment Scores:

Pred. No.: 1,11e-72 Length: 1043
 Score: 758.00 Matches: 140
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 34.33% Indels: 0
 Gaps: 0

US-09-831-452-1 (1-424) x US-10-860-790-19 (1-1043)

QY 285 LeuCySGlyLysLeuMetArgsnGluPheThrIleThrHisValLeuIleProlySgin 304
 Db 1 CTCGTGGAAGAACTGATGAGGAATGAATTATTAACATTACCATGTTCTCATCCCAAGCAA 60
 QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPheLeuIleGlnasp 324
 Db 61 AGTGTGGGTCTGATTATTCGCAACACAGAAAGAAAGAACTTTCTCTCATACAGGAT 120
 QY 325 GlnGluGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrSginThrAlaPheLeu 344
 Db 121 CAGCAGGCTCATACATCAGCTGGCTGATTCATACCAACCCACAGACCCGCTTCTC 180
 QY 345 SerSerValaspIleuHisThrHisCysSerTyrGlnMetLeuProGluSerVala 364
 Db 181 TCCAGTGTGACCTACACACTGCTCTTACAGATGATGTGCCAGAGTCAGTAGCC 240
 QY 365 IleValCysSerProlyspheGluGluThrGlyPhePheLeuThrAspHisGlyLeu 384

Db 241 ATTGTTTCCTCCCAAGTTCAGAACTGATCTTTAACTAAGTACCATGAGCTA 300
 QY 385 GluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerIleAspProProIleu 404
 Db 301 GAGAGATTTCTTCCTGCTCCGCAAGAAAGATTTTCAACACAGAGAGATTCACCTCTG 360
 QY 405 PheCysSerCysSerHisValThrValIleAspArgAlaValThrIleThrAspLeuArg 424
 Db 361 TTCTGTAGCTGCAGCCACGATGACTGTTGTGACAGAGCAGTGAACATCAAGACCTTGA 420

RESULT 10

US-10-767-701-13735
 : Sequence 13735, Application US/10767701
 : GENERAL INFORMATION:
 : APPLICANT: Kovalic, David K.
 : APPLICANT: Zhou, Yihua
 : APPLICANT: Cao, Yonwei
 : TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 : TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 : FILE REFERENCE: 38-21(53535)B
 : CURRENT APPLICATION NUMBER: US/10/767,701
 : CURRENT FILING DATE: 2004-01-29
 : NUMBER OF SEQ ID NOS: 63128
 : SEQ ID NO 13735
 : LENGTH: 1972
 : TYPE: DNA
 : ORGANISM: Sorghum bicolor
 : FEATURE:
 : NAME/KEY: unsure
 : LOCATION: (1)..(1972)
 : OTHER INFORMATION: unsure at all n locations
 : FEATURE:
 : OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS9218_1
 : US-10-767-701-13735

Alignment Scores:

Pred. No.:	4,34e-47	Length:	1972
Score:	529.00	Matches:	163
Percent Similarity:	45.84%	Conservative:	74
Best Local Similarity:	31.53%	Mismatches:	164
Query Match:	23.96%	Indels:	116
DB:	7	Gaps:	22

US-09-831-452-1 (1-424) x US-10-767-701-13735 (1-1972)

QY 10 ProProGluIleAsp---ArgValArgAlaLeuSerGlnLeuGlySerAla-----ValGlu 26
 Db 113 CCACCGCAGCCAGCAGGAGTGGCGCCATCAACATCGAGCGTGGCGCGCCGATCGCC 172
 QY 27 ValAsnGluAspIleProProArgArgTyrPheArgSerGlyValGluIleIleArgMet 46
 Db 173 GTGACCAACCGCATCAGGCTCCCTACTACTTCGCGCATCGCGGCGAGCTCTCCAGAG 232
 QY 47 AlaSerIleTyrSerGlnGluGlnLysAsnIleGlnHisAlaPheIleLeuTyrAsnLysTyr 66
 Db 233 GCTAATATATATCGAAGAGAGATACTTCTCGACCTGATATGATCATCTTCTGAGATAC 292
 QY 67 IleThrLeuPheIleGlnLysLeuProLysHisArgAspTyrLysSerAlaValIlePro 86
 Db 293 TCGAGCTTCTGTGTGACAGCATTCGAGCATGTATTAACATGCTTC-----343
 QY 87 GluLysLysAspThrValIleLysLeuLysGluIleAla---PheProLysAlaGluGlu 105
 Db 344 -----AAGTTAAAGAAAGAGATTTTGTGATTAACCTAATGAT 382
 QY 106 LeuLysAlaGluLeu-----LeuLysArgTyrThr---LysGluTyrThrGluTyrAsn 122
 Db 383 GTTATGAAGAGCTTGACATTTGAGCAAGTGTGACGCGAGGTGGCGAGCATTAAC 442
 QY 123 GluGluLysLysGluAla-----GluGluLeuAlaArgAsnMetAlaIleGlnGln 140
 Db 443 AGAGAGGTACTGTGGAAATCGAATTAATAGTCTTAATGAAATGAACTATGATTAAGTCTGT 502

QY 141 GluLeuGluLysGlnLysArgValAlaGlnGlnLysGlnGln-----156
 Db 503 AGATATAG-----CAGCGCATCCCAACCTCTTATATCTCCAGACCATTTGAGGC 553
 QY 157 -----LeuGlnGlnGlnPheHisAlaPheGlnMetIleArgAsnGln 172
 Db 554 AGCATTAATGAGCATTTGCAAAACCTTCATCTCTGGAGACAGAGGGCATTAATTCGA 613
 QY 173 GluLeuGluLysGlnLysLeuLysIle-----ValGlnLysPhe 185
 Db 614 AGCGTCAGAAACATTAATTAATCAATATCAATAAGAAAGAAACATTAAGTACAC 673
 QY 186 GlyLysValAspPro---GlyLeuGlyGly-----194
 Db 674 TCCATATTAGACCTAATGTCTTAATGGCCGATGGAATGGGCTGTACTGAAATTAAG 733
 QY 195 -----ProLeuValPro---198
 Db 734 GTTCAGTATCCAGCAATTTGAATTAAGACAGAGGATTAACAGATTAAGTATGATCCATCC 793
 QY 199 -----AspLeuGlnLysProSerLeuAspValPheProThrLeuThrVal 213
 Db 794 ATCTTGAACCAAGATGATTCGATGTCCAGTACCTGACCTGCACCTCCACCTCCAGAT 853
 QY 214 SerSerIleGlnProSerAspCysHisThrThrVal-----225
 Db 854 AGCTCTCAAAATGACAAATGACATATGAAATCTGTCTCTCTCTGATGATGTGATGG 913
 QY 225 -----225
 Db 914 TCTGTGACAGAGAAACGCACTCCACTTCTCTGCTAGCTTGAAGAAGATTTTCAA 973
 QY 226 -----ArgProAlaLysProProVal-----ValAspArgSerLeuLys 238
 Db 974 TTGATATCAAAACACCTTCTCTCCACCAAGTTCTGGCAGAGGTGCAAAAGCAATTTCT 1033
 QY 239 ProGluAlaLeuSerAsn---SerGluSerIleProThrIleAspGlyLeuArg-----255
 Db 1034 CCATCTAGAGTTGCTGATCCAAACCCGACCTCCGACCTCAGAAACGCTTTTCAG 1093
 QY 256 HisValValValProGluArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAspThr 275
 Db 1094 AACTTGCAATGTTCATCAAGTATGATGAGTGTCTTCAAGGTGTCGAGTCAAAACCC 1153
 QY 276 AlaArgGlyValIleThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThr 295
 Db 1154 AAAAGAGCTTGAAGAACTGTGGGCTTCTGTCTGTAACCTTGAAGAAAGCACTTTTGC 1213
 QY 296 IleThrHisValLeuIleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsn 315
 Db 1214 GTGACACCTTAAATTAATTCAAAGCAAGAAATGACATTAATAGCGTGAAGCTATAAT 1273
 QY 316 GluGluGluLeuPheLeuIleGlnAspGlnGlnLysLeuIleThrLeuGlyTyrIleHis 335
 Db 1274 GAAAGAGCTTATTTGAAGTTGAGACACAGGCTACCTTCACTCTTGTTGATTCAT 1333
 QY 336 ThrHisProThrGlnThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyr 355
 Db 1334 ACAACATCAACACAGACGCTGCTTCTTCTTCATTAACCTCCACAATCATTAATCTTAT 1393
 QY 356 GlnMetLeuLeuProGluSerValAlaIleValCysSerPro-----LysPheGlnLys 373
 Db 1394 CAGGTCATGCTACCGAAGCAATTCGATAGTTATGAGCACTTACTGACACAAAGAA 1453
 QY 374 ThrGlyPhePheLysLeuThrAsp---HisGlyLeuGlnGlnLysSerSerCysArgGln 392
 Db 1454 CATGTATATTTCATCTCAAGATCCAGGTGATAGGCGATCATGATGTCAGAGAG 1513
 QY 393 LysGlyPheHisProHisSerLysAspPro-----ProLeuPheCysSerCys 408
 Db 1514 AGAGGTTTCATCCATCCAT---AAGCACCTCTTGATGTTGCGCATCTATGAGCAGTGC 1570
 QY 409 SerHisValThrVal---ValAspArgAlaValThrIleThrAspLeuArg 424

```

Db      1571 TCCCATGTGATCATGACACAGATATTAAGTTGATATGATGATCTCGA 1621
|||||
RESULT 11
US-10-425-115-126539
; Sequence 126539, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 126539
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46872C.1
US-10-425-115-126539

Alignment Scores:
Pred. No.: 4,4e-46 Length: 2559
Score: 521.50 Matches: 152
Percent Similarity: 45.02% Conservative: 74
Best Local Similarity: 30.28% Mismatches: 157
Query Match: 23.62% Indels: 119
DB: 6 Gaps: 20

US-09-831-452-1 (1-424) x US-10-425-115-126539 (1-2559)

Qy      25 ValGluValaIaenGluAsePleProBproAargArGtyrPheAargSerGlyValGluIle 44
Db      545 ATCCGCGTGCACCCGATCATGCTCCGCTACTCTCCGATCGCGGCGAGCCTCTC 604
|||
Qy      45 ArgMetAlaSerIleYrSerGluGluGlyAsnIleGluHisAlaPheIleLeuYrAsn 64
Db      605 CGACAGCGCTATATATATCGAGACGAGTAACTTCCGACCTGTCATGTCATCTCTC 664
|||
Qy      65 LysYrIleThrLeuPheIleGluLeuLeuProLysHisArgAspYrLysSerAlaVal 84
Db      665 AGATACTCGAGCTGCTGTGCGAGACGATTCGAGGATCTGTGATTCATGCTTCAG 724
|||
Qy      85 IleProGluYrLys-----AspThrValLysLysLeuYrGluIle 98
Db      725 TTAAGAGAAAGGATTTTGTGATTAACCTGTTATATATCAAGAGCTT----- 775
|||
Qy      99 AlaPheProYrAlaGluGluLeuYrAlaGluLeuLeuYrArgYrThrLysGluYr 118
Db      776 -----GAGACATGAAG---CAAGTTGTGACGCGCGAGATTGTTGACAT 817
|||
Qy      119 -----ThrGluYrAsnGluGluLysLysGluAlaGluGluLeuAla 133
Db      818 AACGAGAGAGGTACTGTGGAATTAT-----ACTAATAGCTTAAT 859
|||
Qy      134 ArgAsnMetAlaIleGlnGlnGluLeuYr-----GluYrGln 147
Db      860 GGAACATATGTACACTCATAGACGAGAGGTACTCCAGCTTATATCTCCACG 919
|||
Qy      148 ArgValAlaGlnGlnYrGlnGlnGlnGluGluGlnGluGlnPheHisAlaPheGluGlu 167
Db      920 ACATTTTACGACAGACAGTAAATGAGCATTG---CAAAACCTTCCATGCTGAGAGACA 976
|||
Qy      168 MetIleArgAsnGlnGluLeuGluYrGluArgLeuYrSile----- 181
Db      977 GTGCGATCATTTATGAGATGTACGAAACATTAATGATATCTACCATATCCAAAGAGAA 1036
|||
Qy      182 ---ValGlnGluPheGlyLysValAspPro---GlyLeuGly 194
Db      182 ---ValGlnGluPheGlyLysValAspPro---GlyLeuGly 194
|||
Db      1037 ACATAGCTAGACACTCCATATTAAGACCTAATGCTTATATGCGCGGTGAATGGCCT 1096
Qy      194 ----- 194
Db      1097 GTTACTGGAATCAAGTTCAATATCCNAGCAATTTGATTAACAAATAATGATATACA 1156
Qy      195 -----ProLeu 196
Db      1157 AGTTAGTCCATCCATCTTGAACCAAGATGATTTGATGCTTCAATGACCTGCA 1216
Qy      197 ValProAspLeuGluYrProSerLeuAspValPheProThrLeuThrValSerSile 216
Db      1217 CCTCACCTCCAGATGATGCTCTCAATATGACATGACATATGAATCTGTTCTCTCTT 1276
Qy      217 Gln-----ProSerAspCysHisThrVal----- 225
Db      1277 GATGATGTTGATGATGCTGTGATCCAGACAGAAACACACTTCTCTTCTGTAATTG 1336
Qy      226 -----ArgProAlaLysProProVal----- 232
Db      1337 GAAGAGAGTTTCTCAATTTGAATATCAACACACTTCTCTTCCACAGTTCTTGACAG 1396
Qy      233 ValAspArgSerLeuYrProGluYrAlaLeuSerAsn---SerGluSerIleProThrIle 251
Db      1397 GTCAAAAGACCAATTTCTCCATCAAGAGTTGCGATCCAAACACAGACTTCCGACTCA 1456
Qy      252 AspGlyLeuArg-----HisValValProGluYrArgLeuYrProGlnPheLeuGln 269
Db      1457 GGAATCGCCGATTTGAGACTGTGATGTCGATCAAGATTAATGAGTGTTCCTTCAAG 1516
Qy      270 LeuAlaSerAlaAsnThrAlaArgGlyValGluThrCysGlyIleLeuYrCysGlyLysLeu 289
Db      1517 GTTGCTGATGCAAAACCAAAAGAGCTTAGAAACCTGTGGGGTCTTCTGATACCTG 1576
Qy      290 MetArgAsnGluPheThrIleThrHisValLeuIleProYrGlnSerAlaGlySerAsp 309
Db      1577 AAAAATGAACCTTTTATGTGACAACTTAATTTCCAAACAGAAATGACATCTGTT 1636
Qy      310 TyrCysAsnThrGluAsnGluGluLeuPheLeuIleGluAspGlnGlnGluIle 329
Db      1637 ACCTGGAAGCTNCAATATGAAGAGACTATTTGAATTCAGACATGGGCTCACTGTT 1696
Qy      330 ThrLeuGlyTyrIleHisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeu 349
Db      1697 ACTCTGTGTTGATTCATACATCCACACAGCTCCTGCTCTTCTCCGTTGACCTC 1756
Qy      350 HisThrHisCysSerTyrGlnMetMetLeuProGluSerValAlaIleValCysSerPro 369
Db      1757 CACATCATTTATTTATATAGGTCATGCTGCTGAAGCATTTGCAATATGATTTGACCT 1816
Qy      370 ---LysPheGlnGluThrGlyPhePheYrLeuThrAsp---HisGlyLeuGluGluIle 387
Db      1817 ACTGACACAAAGAAACATGATATTTATCTCAGAGATCCAGGTGATGGCGCTGATC 1876
Qy      388 SerSerCysArgGlnLysGlyPheHisProHisSerLysAspPro-----Pro 403
Db      1877 CATGATTTGACAGAGAGAGGTTTCCATCTCAT---AAGGACCACTTGATGTTGCGCA 1933
Qy      404 LeuPheCysSerCysSerHisValThrVal---ValAspArgAlaValThrIleThrAsp 422
Db      1934 ATCTNACAGCATGCTCCATGTGATGATGACACCGACCATATAAGTTGATGATGAT 1993
Qy      423 LeuArg 424
Db      1994 CTCGCA 1999
|||
RESULT 12
US-10-425-115-13930/c
; Sequence 13930, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

```

APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants
 FILE REFERENCE: 38-21(53222)B
 CURRENT APPLICATION NUMBER: US/10/425,115
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ. ID NOS: 369326
 SEQ. ID NO 13930
 LENGTH: 1660
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1)..(1660)
 OTHER INFORMATION: unsure at all n locations
 FEATURE:
 OTHER INFORMATION: Clone ID: MRT4577_112696C.1
 US-10-425-115-13930
 Alignment Scores:
 Pred. No.: 8,35e-44 Length: 1660
 Score: 498.00 Matches: 146
 Percent Similarity: 43.51% Conservative: 55
 Best Local Similarity: 31.60% Mismatches: 162
 Query Match: 22.55% Indels: 100
 Gaps: 12
 US-09-831-452-1 (1-424) X US-10-425-115-13930 (1-1660)
 QY 10 ProProGlu-----AsparGValArgAlaLeuSerGln 20
 DB 1461 CCACCTAGATGGCTTCGGAACCTATCTCAGCAGCTAGAGAGACGATGATCGATCGATCC 1402
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgGlyPheArgSerGly 40
 DB 1401 ATGGCGGGGCGCTATGCGCGTGGACCGCATAGCATTCACATCTATTCGCGATCGCT 1342
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluValAsnIleGluHisAlaPhe 60
 DB 1341 GACAACTCCCTCCCGCAGGCTAATGTGTATCGAAGAAAGAAAGATGTTCTGACCTCTAC 1282
 QY 61 IleLeuTyrAsnIleTyrIleThrLeuPheIleGluGluSerProLysHisArgAspTyr 80
 DB 1281 ATCTCTCTCCGATATCTCGAGCTTGTGTGTAGACCATACCCAAACATCCGATTTAT 1222
 QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuGluIleAlaPhe 100
 DB 1221 AATGATTTAAGCTAGGAAAG----- 1198
 QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu 120
 DB 1197 -----GAATTCCTGAAGAA-----GGCCCA 1177
 QY 121 TyrAsnGluGluLysLysGluAlaGluGluLeuAlaArgAsnMetAlaIleGlnGln 140
 DB 1176 CACAACCTGAGAAA-----ATTCTGTGATTGTCAAT 1144
 QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGln-----LysGlnGlnGlnLeu 157
 DB 1143 GAGCTAAGTCCCTTAAGCCAGTGTGAGCAACAATTGGCCACCGGTAATAGCAGGGGT 1084
 QY 158 GluGlnGluGlnPheHisAlaPheGluGluMetIleArgAsnGlnGluGluGluGlu 177
 DB 1083 GCAGATGAACCTAATGCTGTATATGAACCATATCTCGAAGTATGAGTGGAGACACAT 1024
 QY 178 ArgLeuLysIleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeuVal 197
 DB 1023 -----ACCCAGGCCCTTATGTG 1006
 QY 198 ProAspLeuGluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGln 217
 DB 1005 -----TGGAAGTCTTGTCTGGAAGCCCTTACTAAGTACGACGAGAAACCATTA 958

QY 218 ProSerAspCysHisThrValArgProAlaLysProProValAlaAspArgSerLeu 237
 DB 957 CCTGGTTCAAGCATCAACAGCAGCTATGTCGAGATGTTCAACTTATAGACATGCTATG 898
 QY 238 LysProGly----- 240
 DB 897 AAACCATCTACCAACATGCTGTATCCAAAGAGAAACATTATGACACTGATACTA 838
 QY 241 -----AlaLeuSerAsn-SerGluSerIleProThr 250
 DB 837 GGACCTAATGGCTTCAATGATATGAGATGATGATCTTCTAGCCACTAGCAGCATGATG 778
 QY 250 rIleAspGlyLeuArgHisValValProGlyArgLeuCysProGlnPheLeuGlnLe 270
 DB 777 TATTGAGAGTTGAAGATCTATGAAACCCAGTGGCATTTATGAGAGCTTTCCGAGGCT 718
 QY 270 uAlaSerAlaAsnThrAlaArgGlyValGluThrCysGlyIleLeuCysGlyLysLeu 290
 DB 717 CGCTAGATTAATATCCGMAAGATTTGGAAC-TGTGAATCTTCTGCTGATCTGAA 659
 QY 290 LArgAsnGluPheThrIleThrHisValLeuIleProLysGlnSerAlaGlySerAspTyr 310
 DB 658 AAAGAGACCTTTATATGACACATGATGATATTCCTTAAGCAGAAATCAACATCTGATTC 599
 QY 310 rCysAsnThrGluAsnGluGluGluLeuPheLeuIleGlnAspGlnGlnLeuIleThr 330
 DB 598 ATGTAACTAGCATGATGAAGAAGAAATATTGAAGTTCAGACAAAGGCTCCCTTTATC 539
 QY 330 rLeuGlyTyrIleHisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeuHis 350
 DB 538 TCTTGTTGATGATCATACATCACTCAACACACACACCTGCTCTTCACTTCCATGATCTCA 479
 QY 350 sThrHisCysSerTyrGlnMetMetLeuProGluSerValAlaIleValCysSerPro-- 369
 DB 478 TAAATATTAAGATATTCAGATTAATGCTTCGGAAGCAGATTCATATGATGAGCTCTAC 419
 QY 370 ----LysPheGlnGluThrGlyPhePheLysLeuThrAspHis--GlyLeuGluGlu 387
 DB 418 GGATCAACAAGAAACATGCGCATATTCATCTAAGCAGCAGCATGTCAGCGCGTAT 359
 QY 387 eSerSerCysArgGlnLysGlyPheHisProHisSerLys-----AspProProLe 404
 DB 358 CCATGATTTGTATGACATGATGATTCACCTCATGAGAGGCCCTAATGAGAACCAAT 299
 QY 404 uPheCysSerCysSerHisValThrValAlaAspArgAla--ValThrIleThrAspLe 423
 DB 298 CTATGAGCATTCCTCCCATGTATCATGACCAACCAATGTCAAGTTGAGATGACAGACT 239
 QY 423 uArg 424
 DB 238 ACGT 235
 RESULT 13
 US-09-404-520B-5634
 ; Sequence 5634, Application US/09404520B
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Timberlake, William E.
 ; APPLICANT: McClinch, James
 ; APPLICANT: Xu, Jaehyuk
 ; APPLICANT: Gnodski, Azita
 ; TITLE OF INVENTION: Emericella nidulans Genome Sequence and Uses Thereof
 ; FILE REFERENCE: 38-10(15498)A
 ; CURRENT APPLICATION NUMBER: US/09/404,520B
 ; CURRENT FILING DATE: 1999-09-23
 ; NUMBER OF SEQ. ID NOS: 44345
 ; SEQ. ID NO 5634
 ; LENGTH: 8639
 ; TYPE: DNA
 ; ORGANISM: Emericella nidulans
 US-09-404-520B-5634

Alignment Scores:

Prod. No.:	Score:	Length:	Matches:
4.39e-39	465.50	8639	150
Percent Similarity:	42.10%	Conservative:	71
Best Local Similarity:	28.57%	Mismatches:	173
Query Match:	21.08%	Indels:	131
DB:	5	Gaps:	17

US-09-831-452-1 (1-424) x US-09-404-520B-5634 (1-8639)

```

QY      5 GtlyaspValSerleuProProgluAspArgValArgalaLeuSerGlnLeuGlySerAla 24
DB      6475 GGAGCGCGCTCCGACCT-----CAAGTGGGAGAGATCAGCGCTTTCAGCTCAGAGAC 6528

QY      25 ValGluValAspGluAspIleProProArgArgTyrPheArgSerGlyValGluIle--- 43
DB      6529 TACAAATTCACACCTCGCAGATTCCTTACAGTACTGGCTGAGAGCGGCTGCTACCTAGTT 6588

QY      44 -----IleArgMetAla 47
DB      6589 CGTGAGTTAGTAGATGCTCTATGATCTTTGTCACTGTGCTATAGATGATCAGCTTAGGC 6648

QY      48 Ser-IleTyrSerGlnGluGlyAsnIleGluHisAlaPheIleLeuTyrAsnLysTyrI 67
DB      6649 TCGAATCTACCTCGCGGAGAGACATGATGAACAGGCATACCTCTCTTCCTCCGACATGC 6708

QY      67 ethrleupheIleGluLysLeuProLysHisArgAspTyrLysSerAla-----ValI 85
DB      6709 TCAATTAGCTGTCTGTCATTTTAGCCGACATCCGACTTGAAGATGGGAAAAGATGTAG 6768

QY      85 eProGluLysLysAspThrValLysLysLeuLysGlnIleAla-----PheProLysAl 103
DB      6769 GGAAGAACGGAAGCTCTAATAGCGCGCTGAGAGAGAGGTGAACAGAAATCGAGATCCT 6828

QY      103 aGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGluTyrAsnG 123
DB      6829 GGAHAAATGTAAAGCCGCGCATCAACAAGCGATATGACCGATACACAGCTGATGCGTGA 6888

QY      123 uGluLysLysLysGlnAlaGluGluLeuAlaArgAsnMetAla----- 137
DB      6889 ACCGCAATCCGCGCATGCTCTGCGCGCGCTGACGTCCGCTCCGACGCCAGCCGCTC 6948

QY      138 -----IleGlnGlnGluLeuLysGlnLysGlnLysGlnLys 148
DB      6949 GCCACCTCCGAGCCCGCTCTGCTGTGCTGTGTCGGAACCTCTGAGACCTGGAAGCATAA 7008

QY      148 gValAlaGlnGlnLysGlnGlnGlnLysGlnGlnGln----- 161
DB      7009 TGACCTAGCTGTCTCAAGCTGGCTCGGACCGAGATTAGTCGTAGGCAACCTCAGAAACG 7068

QY      162 -----Ph 162
DB      7069 GATGCCCAAGCTGAGCTCAGCGGTGAGATGATTCAGCCGACGCGCGCAGAGATATG 7128

QY      162 eHisAlaPheGlnGluMetIleArgAsnGlnGlu----- 173
DB      7129 GGGGATGTGGAAAGAAAGCTCTGACAGCGAAATGAGAGAGGAGATCCGACGATGTGAGTCG 7188

QY      174 -----LeuGluLysGluArgLeu-----LysI 181
DB      7189 ACCGATTCAGAAAGTGGCACTGAACCTGAGGGGAGCCACAGATTTCAMAGCAAGCTGT 7248

QY      181 eValGlnGlnPheGlyLysValAsp-----ProGlyLeuGlyProLeuValPro-- 198
DB      7249 AAGCGCGCGCTGACCAATCTTCATCTACGCCCACTTACAAATATCCAGACTCTCTCTG 7308

QY      199 -----AspLeuGluLysProSerLeuAspValPheProThrLeuThr--ValS 214
DB      7309 GCAGAAAGCGCTTGTATGTTCCCTCTGTTCGCGTGATATTCACACAGACCAAGCCCGCGCA 7368

QY      214 eIser-----IleGlnProSerAla 220
DB      7369 GCTCACCTAGCAACCGGTTCTCTCCGAAAGAGAGCTTTCATAGTACGACCAAGCG 7428

```

```

QY      220 spCyshIsthrThrValArg----- 226
DB      7429 GGCTTAGCTTCCCGCCGCTCCGAGACAGATTTCACAGACAGTTCGCCGACATCTGCAG 7488

QY      227 -----ProAlaLysProProValValAspArgSer-----LeuLysProGlyAla 242
DB      7489 CAGCAGTGTCTTGAAGTGCAGGGGTCCAGCAAAAGCAAGATTTAGACCCCTCCAGCT 7548

QY      242 euserAsnSerGlnSerIleProThrIleAspGlyLeuArgHis-----ValValValP 260
DB      7549 TCACATTACAGCATCCGCCGACCTTGGAAACGTAAGCTCTCATTCGCCGAGTCTTCTTC 7608

QY      260 roGluArgLeuCySProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValG 280
DB      7609 CACCACAGATGCGGTGACATTTCTGTCTGCTCCGCCGCTCAAAATCTCGTAAGAAATTTGG 7668

QY      280 IuthrCySGlyIleLeuCySGlyLysLeuMetArgAsnGlnuphetrrIleThrHisVal 300
DB      7669 AAACGAGCGGTATTTCTGTGTGACATGATCTAAGCCCTTTATCTCTGACCTTC 7728

QY      300 euIleProLysGlnSerAlaGlySerAspTyrCysAsnThrGlnAsnGlnGluGluLeuP 320
DB      7729 TTAATCCCGAGCAGATATCTACATCAGACACCTGCGAGACAGTCAACGAGTCCGCCATT 7788

QY      320 heLeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTrrIleIsthrHisProThG 340
DB      7789 TTAATTAATCTGCGACTCGGAGATCTAATAGTCTTGCTGGATCCATACGACCCCTAAC 7848

QY      340 IInThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuP 360
DB      7849 AGACGCTTCTTCAAGCTCAAGGACTTTCACACATTCGCGGATTCAGATCATGCTAC 7908

QY      360 roGluSerValAlaIleValCysSerPro--LysPheGlnGlnIuthrGlyPhePheLys 379
DB      7909 CGGAGGCAATGCGATCGTTTCGCGCCGAGCAGACCCCAAACTGGGGGTGATTCGCT 7968

QY      379 euthrAsp---HisGlyLeuGlnGluIleSerSerCysArgGlnLysGly---PheHisP 397
DB      7969 TCAGGACCTCCAGAGCTCAAGTGTGCTGAGCTGTACTGACCGGAGCTTTCATC 8028

QY      397 roHisSerLys 400
DB      8029 CTCATGCTAG 8039

RESULT 14
US-10-779-543-3954
; Sequence 3954, Application US/10779543
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; FILE REFERENCE: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; CURRENT APPLICATION NUMBER: US/10/779,543
; PRIOR FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24

```

PRIOR APPLICATION NUMBER: 60/080,114
PRIOR FILING DATE: 1998-03-31
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 23767
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3954
LENGTH: 730
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4, 9, 22, 364, 377, 505, 512, 602, 625, 626, 655, 667, 684,
LOCATION: 705, 722, 730
OTHER INFORMATION: n = A,T,C or G
US-10-779-543-3954
Alignment Scores:
Pred. No.: 1,47e-38 Length: 730
Score: 445.00 Matches: 109
Percent Similarity: 62.17% Conservative: 34
Best Local Similarity: 47.39% Mismatches: 65
Query Match: 20.15% Indels: 22
DB: 7 Gaps: 3
US-09-831-452-1 (1-424) x US-10-779-543-3954 (1-730)
QY 165 PheGluGluMetLeuLeuArgGngGluLeuGluLysGluArgLeuLysIleValGlnGlu 184
DB 65 TTGAGATGATCACTCAAGAGCAAGAGTACCCGAGGTCAATGCGAATCAGCAAAACC 124
QY 185 PheGlyLysValAlaProGlyLeuGlyGlyProLeuValProAspLeuGlyLysProSer 204
DB 125 TCAGGG-----CTGCAGACGAGATTGATGGAGCGCT 157
QY 205 LeuAspValPheProThr-----LeuThrValSerSerIleGlnPro 218
DB 158 TTGCGCGCTTTTCCACACACGACCAATTCCTGCTGATGATGATTCAGATCAACCT 217
QY 219 SerAspCysHisThrThrValArgProAlaLysProValAlaAspArgSerIleLys 238
DB 218 AATAAAGATGATGACCAACCAATTATGCGCACTCTCTCTGTAACAGGCGCTTACG 277
QY 239 ProGlyAlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAla 258
DB 278 CCAGCTGCTACTCTAAGTGGCTTTCAGATTTAGTGTGAGAGCTGCGATGTAAGT 337
QY 259 ValProGlyArgLeuCysProGlnPheLeuGlnIleuAlaSerAlaAsnThrAlaArgGly 278
DB 338 TTGCGAGAGATCTTTGCCAACAATTCTGCACTGCGCANAAATCTAATACAGTAGAGA 397
QY 279 ValGlnThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHis 298
DB 398 ATAGAACTGTGAATACTGTGGAAACTGACACATATGATTTACTATTACCAT 457
QY 299 ValLeuIleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGlu 318
DB 458 GTAATTGCGCAAGAGAGCTGTGCGGAGCAAGACTATTGTGACATGAGAAATGTTNAGAA 517
QY 319 LeuPheLeuIleGlnAspGlnGlnGlyLeuIleThrIleuGlyThrIleHisThrHisP 338
DB 518 TTATTCAGTGTCCAGATCAACATGATCTCTCTCTTCAAGGAGGAGATTCACATC 577
QY 338 roThrGlnThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMe 357
DB 578 CTACTCAAGTGCATTTTATCCANCGTGTATCTTTAAGTCACTGCTGCTGCTTACCT 637
QY 357 tMetLeuProGluSer--ValAlaIleValCysSerProLysPheGlnGlnThrGly 375
DB 638 TATGTTGCCAAGAGCCNAATGGCATTTNTGGCTACCCAAAGACNTAAGACACTGCC 697
QY 376 PhePheLysLeuThrAsp 381
DB 698 CTCTTANGCTTACCAAT 715

RESULT 15
US-10-767-701-8700
Sequence 8700, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yinhua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
PENDING FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 8700
LENGTH: 909
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28WAY03-CLUS52071_1
US-10-767-701-8700
Alignment Scores:
Pred. No.: 1e-34 Length: 909
Score: 411.50 Matches: 87
Percent Similarity: 62.77% Conservative: 31
Best Local Similarity: 46.28% Mismatches: 59
Query Match: 18.64% Indels: 11
DB: 7 Gaps: 5
US-09-831-452-1 (1-424) x US-10-767-701-8700 (1-909)
QY 230 ProproValAlaAspArgSerLeuLysProGlyAlaLeuSerAsnSerGluSerIlePro 249
DB 15 CCTCATCAAGAGTTCGTCGATCCAGACAGAGCA-----CTGCTACTCTGAAACC----- 65
QY 250 ThrIleAspGlyLeuArgHisValValAlaProGlyArgLeuCysProGlnPheLeuGln 269
DB 66 -----GCAGCCTATCAGAAATTTGCGATGTTCCAGTGCATTAATGAGTGTCTTCAACC 119
QY 270 LeuAlaSerAlaAsnThrAlaArgGlyValGlnThrCysGlyIleLeuCysGlyLysLeu 289
DB 120 GTGCTGAGATTAAATCTGCAAAATTAATTTGAAACCTGTGGAATTCCTGCTGTAAGTCTG 179
QY 290 MetArgAsnGluPheThrIleThrHisValLeuIleProLysGlnSerAlaGlySerAsp 309
DB 180 AAAAAGAGACCTTTTACGATGACGATGATTAATTCCTTAAGCAAGAAATCAACAATCTGAT 239
QY 310 TyrCysAsnThrGluAsnGluGluLeuPheLeuIleGlnAspGlnGlnGlyLeuIle 329
DB 240 TCATGTCAAGCTTCAAAATGAAAGAAATTTGAAATTCAGACAAAGAGCTCCCTTTTG 299
QY 330 ThrLeuGlyTyrIleHisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeu 349
DB 300 TCTCTGTGTGATTCATACACATCCAAACAGAGCTCTTCCATCTTCATTCATGATCTT 359
QY 350 HisThrHisCysSerTyrGlnMetLeuProGluSerValAlaIleValCysSerPro 369
DB 360 CATTAATCATTAAGCATATCAGATTAGCTTACAGAGGAGATGCAATGATGATGCTCT 419
QY 370 -----LysPheGlnGlnThrGlyPhePheLysLeuThrAspHis---GlyLeuGlnGlu 386
DB 420 ACGGATACACAGAAAGAAATGAGCATATTTCATCTCAAGAGGACATGAGGATGGCGCTG 479
QY 387 IleSerSerCysArgGlnLysGlyPheHisProHisSerLys-----AspProPro 403
DB 480 ATTCATGATTGATGAGCGACTGATTCACACCTCATGAGAGCCCTTAGATGGACACCA 539
QY 404 LeuPheCysSerCysSerHisVal 411
DB 540 ATCTATGACATGCTCCCATGTA 563

, Tue Aug 24 10:03:51 2004

us-09-831-452-1.p2n.rnpn

Page 13

Search completed: August 23, 2004, 23:10:11
Job time : 496 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 23, 2004, 16:42:46 ; Search time 534 Seconds
(without alignments)
3373.101 Million cell updates/sec

Title: US-09-831-452-1
Perfect score: 2.08
Sequence: 1 MSDHGVSLLPEDRVRLSQ.....FCSCHVTYVRAVITIDLR 424

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 212409041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=Frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09831452/runat.17082004.151158.23630/app.query.fasta_1.583
-DB=Geneseq.29Jan04 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=humand4.csi
-LIST=45 -POCALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09831452.@CGN.1.1_470@runat.17082004.151158.23630 -NCPU=6 -ICPU=3
-NO.MMAP -ARGOQUERY -NEG.SCORES=0 -WAIT -DSBLOC=100 -LONGLOG
-DEV.TIMEOUT=120 -MAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	1910	3	AAAS9852 Human AMS
2	2208	100.0	2003	5	AAFS7478 Human SAP
3	1904	86.2	1384	3	AAAS9853 Murine AM
4	1269.5	57.5	1621	7	ABZ56929 Human AMS
5	1262.5	57.2	1392	7	ABZ56931 Mouse AMS
6	1262.5	57.2	1428	7	ABZ56930 Mouse AMS
7	1236	56.0	1973	6	ABK13102 Human sec
8	1229.5	55.7	4052	4	AAIS9567 Human pol

9	1229.5	55.7	4714	6	ABK92424 Human pro
10	1102.5	49.9	1236	5	AAAS85218 DNA encod
11	920.5	41.7	1318	4	AAI61352 Human pol
12	871.5	39.5	1611	4	ABL03889 Drosophill
13	815.5	36.9	1898	9	ADBL01940 Human cDN
14	758	34.3	1041	3	AAAS9708 Human ova
15	758	34.3	1041	6	ABN72602 Ovarian c
16	758	34.3	1041	8	ADA08767 Human ova
17	758	34.3	1043	8	ADA08769 Human ova
18	758	34.3	1043	6	ABN72603 Ovarian c
19	758	34.3	1043	8	ADN08768 Human ova
20	656	31.5	4331	8	ABL03898 Drosophill
21	654	29.6	407	8	ACH47766 Human inf
22	489	22.1	1877	3	AAAC9460 Arabidops
23	481	21.8	406	8	ACH47349 Human inf
24	472	21.4	348	8	ACH31362 Human bon
25	463	21.0	431	8	ACH47326 Human inf
26	445	20.2	730	2	AAZ15641 Human gen
27	433	19.6	301	4	AAAF92297 Bovine ma
28	387	17.5	277	3	AAAC02840 Human sec
29	385.5	17.5	1025	3	AAAC46416 Arabidops
30	385	17.4	1010	3	AAAC40578 Arabidops
31	321	14.5	775	4	AAH33413 Human col
32	294	13.3	386	4	AAAS8882 Novel hum
33	257.5	11.7	387	5	AAAF7287 Novel hum
34	225.5	10.2	416	4	AAAS8474 Novel hum
35	216.5	9.8	1184	4	AAI06901 Human rep
36	216.5	9.8	1184	4	ABA08120 Human ova
37	188.5	8.5	397	4	AAAS8494 Novel hum
38	146.5	6.6	5889	5	AAAS7563 DNA encod
39	138.5	6.3	2581	4	ABL26445 Drosophill
40	138.5	6.3	2655	4	ABL20571 Drosophill
41	133	6.0	5109	4	ABL17409 Drosophill
42	133	6.0	40152	4	ABL17408 Drosophill
43	132.5	6.0	4969	5	AAAS81653 DNA encod
44	132.5	6.0	10030	4	ABL15505 Drosophill
45	132.5	6.0	20245	4	ABL15504 Drosophill

ALIGNMENTS

RESULT 1	AAAS9852	AAAS9852 standard; cDNA; 1910 BP.
ID	AAAS9852	
XX	AAAS9852	
AC	AAAS9852	
XX		
DT	13-OCT-2000	(first entry)
XX		
DE	Human AMSH nucleotide sequence.	
XX		
KW	AMSH; signal transduction; SH3 domain; cytokine; STMW; severe infection;	
KW	cancer; autoimmune disorder; human; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200029436-A1.	
XX		
PD	25-MAY-2000.	
XX		
PF	12-NOV-1999;	99WC-JP006309.
XX		
PR	12-NOV-1998;	98JP-00322674.
XX		
PA	(NISC-) JAPAN SCT & TECHNOLOGY CORP.	
XX		
PI	Sugamura K, Tanaka N;	
XX		
DR	WPI; 2000-387743/33.	
XX	F-PSPB; AAB03791.	
XX		
PT	Protein AMSH being signal transducer interacting with SH3 domain of	
PT	cytokine signal transducer STMW, for diagnosis and development of drugs	

PT of treating diseases e.g. severe infection, cancers and autoimmune disorders.

XX Claim 4; Page 14-15; 22pp; Japanese.

XX This sequence represents the human AMSH gene. The AMSH protein is involved in signal transduction, and interacts with the SH3 domain of cytochrome signal transducer SHAM. The invention includes human and murine AMSH protein and gene sequences (hAMSH and mAMSH), and also includes antibodies directed against hAMSH and mAMSH. The AMSH protein can be used in the diagnosis and development of drugs for treating diseases due to the dysfunction of cytochrome signal transduction e.g. severe infection, cancers and autoimmune disorders.

XX Sequence 1910 BP; 578 A; 424 C; 421 G; 487 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,086-189	Length:	1910
Score:	2208.00	Matches:	424
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-831-452-1 (1-424) x AAS5852 (1-1910)

QY	1	MetSerAspHisGlyAspValSerLeuProGluAspArgValAlaLeuSerGln	20
DB	11	ATGCTGACCATGGAGATGTGAGCTCCGCCGAGACCGGGTGAGGGCTCTCTCCAG	70
QY	21	LeuGlySerAlaValGluValAsnGluAspIleProProArgArgGlyPheArgSerGly	40
DB	71	CTGGGATGCGGAGAGGTGAAAGACATTCACCCCGTGGACATTCCTCGCTTGA	130
QY	41	ValGluIleIleArgMetAlaSerIleTyrSerGlnGluGlnValAsnIleGluHisAlaPhe	60
DB	131	GTTGAGATTATCCAAATGCGATTCATTACTCTGAGAGAGCAACATTGACACTGCTTC	190
QY	61	IleLeuTyrAsnIleTyrIleThrLeuPheIleGluLeuLeuProIleHisArgAspTyr	80
DB	191	ATCCTTATTAACAAGTATATCATCGCTCTTATTGAGAACTACCAAAACATCGAGATTAC	250
QY	81	LysSerAlaValIleProGluLeuLeuAspThrValIleValLeuLeuGluIleAlaPhe	100
DB	251	AAATCTCTGCTCATCTCTGAAAAGAAAGACAGTAAGAAATTAAGAGATTGCAATT	310
QY	101	ProIleAlaGluGluLeuLeuAlaGluLeuLeuLeuTyrThrIleGluIleThrGlu	120
DB	311	CCCAAGCAGAGAGCTGAGAGCGAGAGCTGTAAAGATATACCAAGAAATATACAGAA	370
QY	121	TyrAsnGluGluLeuLeuLeuGluAlaGluGluLeuAlaArgAsnMetAlaIleGlnGln	140
DB	371	TATATAGAGAAAGAAAGAGAGAGAGAGAAATGGCCCGGAAATGGCCATCCAGCAA	430
QY	141	GluLeuGluLeuGluLeuGluLeuValAlaGluGluLeuGluGluGluGluGlnGlu	160
DB	431	GAGCTGAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	490
QY	161	GlnPheHisAlaPheGluGluMetIleArgAsnGlnIleGluGluLeuGluAlaGluLeu	180
DB	491	CAGTTCATGCGCTCGAGAGATGATCGGAACAGAGAGCTAAGAAAGAGAGAGAGAA	550
QY	181	IleValGlnGluPheGluLeuValAspProGluLeuGlyProLeuValProAspLeu	200
DB	551	ATTGTACAGAGAGTTGGAGAGTGAACCTGGCTTGAAGTGGCCCGCTAGCTTGACTTG	610
QY	201	GluLeuProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp	220
DB	611	GAGAGAGCGCTCTAGATGTGTCCCACTTAACATGCTCATCCATACAGAGCTTCAGAC	670
QY	221	CysHisThrThrValArgProAlaValProProAlaValAspArgSerLeuLeuProGly	240
DB	671	TGTACACAACTGTAAGGCCAGTAAGCCACTGTGTGAGACAGGTCTCTTGAACCTGGA	730

QY	241	AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaValPro	260
DB	731	GCACTGAGCAACTAGAAAGATATCCACATCATGATGGCCCATGTGGGTGCTCCT	790
QY	261	GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu	280
DB	791	GGGCGGCTGTCCCAAGATTTCTCCAGTTAGCCAGTCCCAACATGCGCGGAGTGAG	850
QY	281	ThrCysGlyIleLeuCysGlyLeuLeuMetArgAsnGluPheThrIleThrHisValLeu	300
DB	851	ACATGTGAAATTTCTGTGAAAACCTGATAGAGATTAATTTACATTACCAATGTTCTC	910
QY	301	IleProGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe	320
DB	911	ATCCCAAGCAAGAGCTGGGTGATTACTGCAACAGAGAGAGAGAGAGAACTTTTC	970
QY	321	LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyThrIleHisThrHisProThrGln	340
DB	971	CTCATACAGATCAGAGGCTCATACACTGGGCTGATTTCACTACACCCCAACAG	1030
QY	341	ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro	360
DB	1031	AACGGTTTCTCTCAGATGTGACCTACACACTCAGTCTCTTACAGATGATTTGCCA	1090
QY	361	GluSerValAlaIleValCysSerProIlePheGlnGluThrGlyPhePheLeuThr	380
DB	1091	GAGTCAGAGCAATGTTGTCTCCCAAGTCCAGAGAACTGATTTTAACCTAACT	1150
QY	381	AspHisGlyLeuGluGluIleSerSerCysArgGlnIleGlyPheHisProHisSerIle	400
DB	1151	GNACATGACCTAGAGAGATTTCTTCTGTCGCAAGAAAGATTTCTACACAGCAG	1210
QY	401	AspProProLeuPheCysSerCysSerHisValThrValAlaAspAlaValThrIle	420
DB	1211	GATCCACTGTGTCTGTAGCTGACGACGACGTCGTGTGTGACAGAGAGTGCATC	1270
QY	421	ThrAspLeuArg	424
DB	1271	ACAGACTTCGA	1282

RESULT 2
AAS57478
ID AAS57478 standard; DNA; 2003 BP.
AC AAS57478;
XX 11-JUN-2001 (first entry)
DT XX
DE Human SAPI/AMSH nucleotide sequence.
XX Smad associating polypeptide; SAP; cell cycle; TGF-beta; SAPI/AMSH;
XX signal transduction; phosphorylation; gene therapy; human; ds.
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 63..1337
FT /*tag= a
FT /product= "SAPI/AMSH"
PN WO200121794-A2.
XX
XX 29-MAR-2001.
XX
XX 20-SEP-2000; 2000MO-US025790.
XX
XX 20-SEP-1999; 99US-0154846P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Itoh F, Itoh S, Heldin C, Ten-Dijke P;
PI
XX

DR MPI: 2001-257987/26.
 DR P-PSDB: AAB62193.
 PT Novel Smad associating proteins useful for regulating cell cycle in a
 PT mammalian cell, and for identifying lead compounds for pharmacological
 PT agents.
 XX
 PS Example 1, Page 53-55; 90pp; English.
 CC The invention relates to Smad associating polypeptides (SAP) and nucleic
 CC acids encoding the SAPs. The SAPs can be expressed by standard
 CC recombinant methodology. The SAP proteins and polynucleotides are useful
 CC for regulating the cell cycle in a mammalian cell, and for modulating TGF
 CC -beta superfamily signal transduction in a mammalian cell. SAP
 CC polypeptide are useful for testing Smad binding, for testing
 CC phosphorylation, for generating antibodies, and as components of an
 CC immunoassay. The SAP polynucleotides are useful in gene therapy. The
 CC present sequence represents the nucleotide sequence of human SAP1/AMSH
 CC that can be used in modulating TGF-beta superfamily signal transduction
 CC in a mammalian cell
 XX
 SO Sequence 2003 BP, 626 A; 439 C; 441 G; 497 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,15e-189 Length: 2003
 Score: 2208.00 Matches: 424
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 5
 US-09-831-452-1 (1-424) x AAF57478 (1-2003)

Qy 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
 Db 63 ATGCTGACATGAGATGTAGCCTCCGCCGCAAGACGGAGGAGCTCTCTCCAG 122
 Qy 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 Db 123 CTGGGTGTGGCGTGAAGGTGAATGAAAGACATTCACCCGCTGGTCTCCGTCGGA 182
 Qy 41 ValGluIleIleAsgMetAlaSerIleTyrSerGluGluGluIleGluHisAlaPhe 60
 Db 183 GTTAGAGTATATCCCAATGGCATTCATCTGAGGAAGGCAACATTTGAACTGCTTC 242
 Qy 61 IleuTyrAsnIleTyrIleThrLeuPheIleGluIleuSerLeuProLysHisArgAspTyr 80
 Db 243 ATCCTCTATTAACAGTATATACGCTCTTTTATAGAACTACCAAAACATCGAGATTAC 302
 Qy 81 LysSerAlaValIleProGluIleAspThrValIleLysIleuLysGluIleAlaPhe 100
 Db 303 AAATCTGCTGTCATTCCTGAAAAGAAAGACACAGTAAAGAAATTAAGGAGATTGCTTT 362
 Qy 101 ProLysAlaGluGluIleuLysAlaGluLeuIleuLysArgTyrThrLysGluTyrThrGlu 120
 Db 363 CCCAAAGCAAGAGAGCTGAAGGCGAGCTGTAAAGCAATATACCAAGAAATATACGAA 422
 Qy 121 TyrAsnGluIleuLysIleuLysGluAlaGluIleuAlaArgAsnMetAlaIleGlnGln 140
 Db 423 TATATGTAAAGAAAGAAAGAAAGCAAGAGAAATTTGCCCGGAAACATGGCCATCCAGCA 482
 Qy 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnIleuLysGlnIleuGluGlnGlu 160
 Db 483 GAGCTGAAAAAGAAAAAGAGAGGTAGCAACAAGAGAGCGCATTTGGAACAGGAA 542
 Qy 161 GluPheHisAlaPheGluGluMetIleArgAsnGlnIleuLeuGluLysGluArgLeuLys 180
 Db 543 CAGTTCCATGCTTCGAGAGATGATCCGAAACCAAGAGAGCTAGAAAAAGCACTGAAA 602
 Qy 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeu 200
 Db 603 ATTGTACAGAGGTTGGGAAAGGTAGACCTGGCTAGGTGCGCGCTAGTGACTTG 662

Qy 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 Db 663 GAGAAAGCCCTCTTGAATGTGTTCCTCCACCTTAACAGTTCATCCATACAGCTTCAGAC 722
 Qy 221 CysHISThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
 Db 723 TGTCAACAACAGTGAAGGCGAGCTAAGCCACTGTGTGTGAAGGTCTTGAACCTGSA 782
 Qy 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValAlaPro 260
 Db 783 GCACGTAGCAACCTGAAAGATATCCCAATCAATCATGATTCGCGCATGTGTGGTGCCT 842
 Qy 261 GlyArgLeuCyseProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
 Db 843 GGGGGCTGTGCCACAGTTTCTCCAGTTAGCCAGTCCACACACGCGCGGGAGTGGAG 902
 Qy 281 ThrCyGlyIleLeuCyGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 Db 903 ACAATGGAATTCCTGTGTGAAACCTGATGAGAGATGAATTAACATTAACCATGTTCTC 962
 Qy 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 Db 963 ATCCCAACCAAGTGTGTGTGTGATTCATCTCAACACAGAAAGAAAGAACTTTTC 1022
 Qy 321 LeuIleGluAspGlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
 Db 1023 CTCATACAGATACGAGAGGCTTCATACATCGGCTGATTCATATCACTCCACACAG 1082
 Qy 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGluMetLeuPro 360
 Db 1083 ACCGCTTCTCTCCAGTGTGACCTACACACTCACTGCTCTTACAGATGATGTGGCA 1142
 Qy 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380
 Db 1143 GAGTCAGTACCATGTTGTCTCCCAAGTTCAGAAACCTGGATCTTTAACTAACT 1202
 Qy 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 Db 1203 GACCATGAGCTAGAGAGATTTCTTCCTGTGTCCAGAAAGATTCATCATCACAGCAAG 1262
 Qy 401 AspProProLeuPheCyseSerCysSerHisValThrValAspArgAlaValThrIle 420
 Db 1263 GATCAACCTCTGTTCTGTAGTCGACGACGAGCTGTGTGACAGAGCACTGACCATC 1322
 Qy 421 ThrAspLeuArg 424
 Db 1323 ACAGACCTTCGA 1334

RESULT 3
 AAAS9853
 ID AAAS9853 standard; cDNA; 1384 BP.
 XX
 AC AAAS9853;
 XX
 XX 13-OCT-2000 (first entry)
 DT
 XX
 DE Murine AMSH nucleotide sequence.
 XX
 XX AMSH, signal transduction; SH3 domain; cytokine; STRM; severe infection;
 XX cancer; autoimmune disorder; mouse; ss.
 OS Mus sp.
 XX
 XX WO200029436-A1.
 XX
 XX 25-MAY-2000.
 PD
 XX
 XX 12-NOV-1999; 99WO-JP006309.
 PF
 XX
 XX 12-NOV-1998; 98CP-00322674.
 PR
 XX
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Sugamura K, Tanaka N;
 XX
 DR WPI: 2000-387743/33.
 DR P-PSDB: AAB03792.
 XX
 PT Protein AMSH being signal transducer interacting with SH3 domain of
 PT cytokine signal transducer STRM, for diagnosis and development of drugs
 PT of treating diseases e.g. severe infection, cancers and autoimmune
 disorders.
 XX
 PS Claim 10; Page 18-19; 22pp; Japanese.
 XX
 CC This sequence represents the murine AMSH gene. The AMSH protein is
 CC involved in signal transduction, and interacts with the SH3 domain of
 CC cytokine signal transducer STRM. The invention includes human and murine
 CC AMSH protein and gene sequences (hAMSH and mAMSH), and also includes
 CC antibodies directed against hAMSH and mAMSH. The AMSH protein can be used
 CC in the diagnosis and development of drugs for treating diseases due to
 CC the dysfunction of cytokine signal transduction e.g. severe infection,
 CC cancers and autoimmune disorders
 CC
 SQ Sequence 1384 BP; 402 A; 354 C; 323 G; 305 T; 0 U; 0 Other;
 CC
 Alignment Scores:
 Pred. No.: 2, 14e-162 Length: 1384
 Score: 1904.00 Matches: 354
 Percent Similarity: 91.75% Conservative: 35
 Best Local Similarity: 83.49% Mismatches: 35
 Query Match: 86.23% Indels: 0
 DB: 3 Gaps: 0
 US-09-831-452-1 (1-424) x AAAS9853 (1-1384)
 QY 1 MetSerAspHisGlyAspValSerLeuProProGluAspArgValArgAlaLeuSerGln 20
 DB 56 ATGCTCAACCATGGGATGTGAGCTCCACCCCAACCGGGTGGAGATTCTGTCCCAA 115
 QY 21 LeuGlySerIleValGluValAsnGluAspIleProProArgArgTyrPheArgSerGly 40
 DB 116 CTGGGAGTGCAGTGGTGAATGAAGACATTCACCCCGTGGTACTACCCCTCGGT 175
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGluValAsnIleGluHisAlaPhe 60
 DB 176 GTTAGATCATCCGATGCGCTCGCTTACTCGAAGAAAGCAACTTGAACATGCTTT 235
 QY 61 IleLeuTyrAsnLeuTyrIleThrLeuPheIleGluLeuSerProLysHisArgAspTyr 80
 DB 236 ATCTCTAACACAGTACATCAGCTGTTATTGAAAACTTCGAAACCGAGACTAC 295
 QY 81 LysSerAlaValIleProGluTyrLysAspThrValLysLysLeuLysGluIleAlaPhe 100
 DB 296 AAATCACTATCATCTCTGAGAAAGAAAGATGCTGTCAGAAATTAAGAGCGTCGCTTC 355
 QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrTrpGln 120
 DB 356 CCTAAACCGGAAAGCTGAAGACAGAGCTCTTGAAGAGATACCCAAAGAAATGACGAG 415
 QY 121 TyrAsnGluGluTyrLysGluAlaGluGluLeuAlaArgAsnMetAlaIleGlnGln 140
 DB 416 TATTAAGAGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 475
 QY 141 GluLeuGluTyrGluTyrGluValAlaGluGluLysGluGluGluGluGluGluGlu 160
 DB 476 GAGTTGAGAAAGAAAGAAAGAGAGGTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 535
 QY 161 GlnPheHisAlaPheGluGluGluMetIleArgAsnGluGluGluGluGluGluGlu 180
 DB 536 CAATTCCATGCTTTGAGAGATATCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555
 QY 181 IleValGlnGluPheGluTyrValAspProGlyLeuGlyLysProLeuValProAspLeu 200
 DB 596 ATTGTTCAAGAGTTCGGGAAGGTAGACCTGCGCCCTGCGGGCTGCTGCTCGATCTG 655

QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 DB 656 GAAAGACCTGTGTGATGTGGCCCCCAGCTCAGCTTCTGCCCCAGCAGACTCCAGAC 715
 QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
 DB 716 TGTAAACAGCAGAGAGCCAGCTAAGCCACTGTGTGAGCAGAGTCCCTGAACCTGGA 775
 QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValValPro 260
 DB 776 GCGTTAAGCGTCATAGAAATGTTCACCATTCGAAAGAGCCCTGCGCCCATCGTGGTCCC 835
 QY 261 GlyArgLeuCysProGlnPheLeuGluLeuAlaSerAlaAsnThrAlaArgGlyValGlu 280
 DB 836 CGTATCTGTCTCAGAAATTTCTCCAGCTTCCAGTCCCATACCGCCAAAGCATGGA 895
 QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 DB 896 ACGTGTGAGTCTGTGTGAGAAACTGATGAAATGAAATTCACAAACACACATGTCTC 955
 QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 DB 956 ATCCCAAGACAAATGTGGGCTGTGATTTATTCACACGAGAAATGAAGAAATTTTC 1015
 QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
 DB 1016 TTTATGACAGATGACCTTGACCTCCTCACTTGTGTGATCATCTACCTCAACCCAA 1075
 QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuPro 360
 DB 1076 ACGGCTTCTGTGTGAGTGTGATCTCCACCTACCTGCTTCAACAAATGATTTTCA 1135
 QY 361 GluSerValAlaIleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThr 380
 DB 1136 GAGTCAACGCAATGCTGTCTTCCCAAGTCCAGGAATGATCTTTAACTACT 1195
 QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 DB 1196 GACTATGCTTCAAGAAATTTCAACCTGCGCGAGAAAGCTTTCACCCCAAGGAGCA 1255
 QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspArgAlaValThrIle 420
 DB 1256 GACCCACGCGTGTGTGTGACTGACGACCATCTCTGTCMAAGACAGAAATGTGACGATC 1315
 QY 421 ThrAspLeuArg 424
 DB 1316 ACAAGACTTCGA 1327
 DB
 RESULT 4
 AB256929
 ID AB256929 standard; DNA; 1621 BP.
 AC
 XX AB256929;
 DT 04-APR-2003 (first entry)
 XX
 DE Human AMSH-LP DNA # SEQ ID 1.
 XX
 XX Human; AMSH-LP; neuroprotective; immunomodulator; cell death;
 XX nervous system; immune system; gene; ds.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 34..1344
 FT CDS /tag= a
 FT /product= "AMSH-LP"
 FT
 PN WO2002102144-A1.
 XX
 XX 27-DEC-2002.
 PD
 XX
 PF 17-JUN-2002; 2002WO-JP006003.

XX 18-JUN-2001; 2001JP-00184136.
 PR (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX Sugamura K, Ishii N;
 XX WPI: 2003-167427/16.
 DR P-PSDB; ABP60075.
 XX
 PT Cell death-inducing model non-human animals specifically in nervous or
 PT immune system for study and analysis of cytokine signal transfer
 PT mechanism, applicable in screening promoters or inhibitors.
 XX
 PS Claim 20; Page 42-45; 68pp; Japanese.
 XX
 CC The invention relates to a cell death-inducible model non-human animal,
 CC characterised by deletion of the AMSH or/and AMSH-LP gene function on the
 CC chromosome. The model animals are applicable in screening promoters or
 CC inhibitors to treat and diagnose diseases in the nervous or immune system
 CC caused by deletion of AMSH and/or AMSH-LP. The current sequence
 CC represents a human AMSH-LP gene fragment
 XX
 SO Sequence 1621 BP; 523 A; 317 C; 336 G; 445 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6,32e-105 Length: 1621
 Score: 1269.50 Matches: 240
 Percent Similarity: 73.02% Conservative: 74
 Best Local Similarity: 55.81% Mismatches: 101
 Query Match: 57.50% Indels: 15
 Ds: Gaps: 2

US-09-831-452-1 (1-424) x ABZ56929 (1-1621)

QY 1 MetSerAspHisGluYAspValSerLeuProGluLysPArgValAlaSerGln 20
 Db 79 ATGCGTAAACCATACAGATGTTTCCCTTAAGCCAGAAAGAGAGTCCCTGAGCCTTAAGCAG 138
 QY 21 LeuGlySerAlaValGluValAsnGluAspLeuProGluArgArgIYrPheArgSerGly 40
 Db 139 CTGCTGTGTAATATACCATCAGTGAAGACATCACTCACACGCTTACTTAGGTGGA 198
 QY 41 ValGluLeuLeuLeuMetAlaSerIleTyrSerGluGluGluValLeuIleGluHisAlaPhe 60
 Db 199 GTAGAGATGAGAGAGAGGCGCTGTGTATTGGAAAGAGAAATTTGAAAAATGCTTT 258
 QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 Db 259 GTTCTTATATATAATTATTAACCTTATTGTGAAAGAGTCTCTTAACCATCGAGATTAC 318
 QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLysLeuGluIleAlaPhe 100
 Db 319 CAGCATATGTGACAGTACCTGAAAGCAGATATTATGAAAGAACTGAAAGAGATTGCAATTC 378
 QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGluTyrThrGlu 120
 Db 379 CCAAGGACAGATGATTAAGAAAGAGACCTTTTAAAGAAATTAACGTAGAAATCCAAAGAA 438
 QY 121 TyrAsnGluGluLysLysLysGluAlaGluGluLeuAlaArgAsnMetAlaIleGlnGln 140
 Db 439 TATTTCGAAAGCAAAAACCAATATAAAGCTGAAATCTCAAAAATTTGAGCATCAGAGA 498
 QY 141 GluLeuGluLysGluLysGluArgValAlaGlnGlnLysGlnGlnGlnGlnGlnGln 160
 Db 499 TTGATAGAGGACAGAAAGAGGAGGATGCTCAGTGGCCAGCAGCAGCTAGAAATCGAG 558
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluGluGluGluArgLeuLys 180
 Db 559 CAGTTTCTGTTTTCGAAAGATCAACTCAGAGCAGAAAGTTAGCCGAGGTCAAAATGCGA 618
 QY 181 IleValGlnGluPheGlyLysValAspProGlyLeuGlyLysProLeuValProAspLeu 200

Db 619 AGTCAGCAAAACCTCAGGG-----CTGTCAAGCAGCAT 651
 QY 201 GluLysProSerLeuAspValPheProThr-----LeuThrValSer 214
 Db 652 GATGGAGGCGTTTGTCTGCTTTCCACACACCAAGCAATTCCTTGCTGATGATATT 711
 QY 215 SerIleGlnProSerAspCysHisThrThrValArgProAlaLysProValValaP 234
 Db 712 GCGATCAACCTAATAAAAGTATGATCAACCAATTATGTCAGCAGCATCTCTCTGTAAC 771
 QY 235 ArgSerLeuLysProGlyAlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeu 254
 Db 772 AGGCGCTTAACCCAGCTGCTCTAAGTGTGTTCAGAAATTAGTGTTAAGAGACTG 831
 QY 255 ArgHisValValValProGluLysGluCysProGlnPheLeuGluLeuAlaSerAlaAsn 274
 Db 832 CGATGTGTGTTTGTCCAGAAAGATCTTCCCAAAATTTCTCAACTGGCAGAAATCTAAT 891
 QY 275 ThrAlaArgGlyValGluThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPhe 294
 Db 892 ACAGTGAAGAGATGAAGAACTGTGGAATACTGTGCAAAACTGACACATATATGATTT 951
 QY 295 ThrIleThrHisValLeuLeuLeuProLysGlnSerAlaGlySerAspTyrCysAsnThrGlu 314
 Db 952 ACTATTAACCATGATATGTGTGCCAAAGCAGTGTGGGACAGACTATTGTGACATGGAG 1011
 QY 315 AsnGluGluGluLeuPheLeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrPle 334
 Db 1012 AATGTAGAGGATTTATTCATGTTTCAAGATCAACATGATCTCTCACTTAGAGATGAT 1071
 QY 335 HisThrHisProThrGlnThrAlaPheLeuSerValAspLeuHisThrHisCysSer 354
 Db 1072 CATACACATCCACCACTCAACTGATTTATACAGCGTGTGATCTTCACATCACTACGTCTCC 1131
 QY 355 TyrGlnMetMetLeuProGluSerValAlaIleValCysSerProLysPheGlnGluThr 374
 Db 1132 TATCAACTATATTGTCAGAGGCACTTGCCATTTGTTGCTCCAAAGCATMAAGACT 1191
 QY 375 GlyPhePheLysLeuThrAspHisGlyLeuGluGluIleSerSerCysArgGlnGly 394
 Db 1192 GGCATCTCAGGCTCAGCAATCTGTCATGCTTGAAGTTTCTGTGTTGTAATAAAGGCG 1251
 QY 395 PheHisProHisSerLysAspProProLeuPheCysSerCysSerHisValThrVal 414
 Db 1252 TTTCAATCCACACCAAGAGGCCAGGCTGTCACTATATGCAAAACATGTCTTGTAATA 1311
 QY 415 AspArgAlaValThrIleThrAspLeuArg 424
 Db 1312 GACATMAAAATATATGTTGTGATCTGAGG 1341

RESULT 5
 ABZ56931
 ID ABZ56931 standard; DNA; 1392 BP.
 XX
 AC ABZ56931;
 XX
 DT 04-APR-2003 (first entry)
 XX
 XX Mouse AMSH-LP DNA # SEQ ID 5.
 DE
 XX Mouse; AMSH-LP; neuroprotective; immunomodulator; cell death;
 KW nervous system; immune system; gene; ds.
 XX
 OS Mus musculus.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 72..1382
 FT /*tag= a
 FT /product= "AMSH-LP"
 XX
 PN MO2002102144-A1.
 XX
 PD 27-DEC-2002.

XX 27-DEC-2002.
 PD 17-JUN-2002; 2002MO-JP006003.
 PF 18-JUN-2001; 2001JP-00184136.
 PR (MISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX Sugamura K, Ishii N;
 XX WPI: 2003-167427/16.
 DR P-PSDB; ABP60076.
 XX
 PT Cell death-inducing model non-human animals specifically in nervous or
 PT immune system for study and analysis of cytokine signal transfer
 PT mechanism, applicable in screening promoters or inhibitors.
 XX
 PS Claim 23; Page 48-51; 68pp; Japanese.
 CC The invention relates to a cell death-inducible model non-human animal,
 CC characterised by deletion of the AMSH or/and AMSH-LP gene function on the
 CC chromosome. The model animals are applicable in screening promoters or
 CC inhibitors to treat and diagnose diseases in the nervous or immune system
 CC caused by deletion of AMSH and/or AMSH-LP. The current sequence
 CC represents a mouse AMSH-LP gene fragment
 XX
 XX Sequence 1428 BP; 412 A; 336 C; 345 G; 335 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 2.29e-104 Length: 1428
 Score: 1262.50 Matches: 241
 Percent Similarity: 72.58% Conservative: 74
 Best Local Similarity: 55.53% Mismatches: 96
 Query Match: 57.18% Indels: 23
 Gaps: 4
 US-09-831-452-1 (1-424) x ABZ56930 (1-1428)

DB 633 CAGTTCCTGTTTTTGGAGCACTCAAGAGCAGAGTTGGCTCGAGCCAGATTGCA 692
 QY 181 ILeValGInGluPheGlyLySValAspProGlyLeuGlyGlyProLeuValProAsp--- 199
 DB 693 -----GGTTCAGGACTCTCCGGTGTGTCTAGAGC 722
 QY 200 LeuGluLySProSerLeuAspValPheProThrLeuThrValSerSerTLeuInProSer 219
 DB 723 ACTGACGGAGAGTGCCTGCTCTTTCACCCACAGACCACTCTGAGAAACGA 782
 QY 220 AspCysHisThrThrValArgProAlaLys-----Pro 230
 DB 783 TTTCAGAT-----CACCTCATAAATGATGAGAGCAATTGGCTACTACTCT 833
 QY 231 ProValValAspArgSerLeuLySProGlyAlaLeuSerAsnSerGluSerTLeuProThr 250
 DB 834 CTTCCCGTGAACAGGCGCTTAAGCCAGCAGCCACTGAGCTGTTCCAGATTGGTG 893
 QY 251 IleAspGlyLeuArgHisValValValProGlyArgLeuCysProGlnPheLeuGlnLeu 270
 DB 894 GTTGAAGGCTGAGGCTGTAGTATTATCAAGATCTTTGCCATPAAATTCTGCTCG 953
 QY 271 AlAserAlaAsnThrAlaArgGlyValGluThrCysGlyTLeuCysGlyLysLeuMet 290
 DB 954 GCTGACTTCAACAGGTGAGAGATAGAGACTGTGGGATCTCTGTGGAAAACTGACA 1013
 QY 291 ArgAsnGluPheThrTleThrHisValLeuIleProLySglnSerAlaGlySerTyr 310
 DB 1014 CACAATGATTCACCATATCATGATGTGTGCCAAAGCACTGTGGCCCAACATAT 1073
 QY 311 CysAsnThrGluAsnGluGluLeuPheLeuIleGlnAspGlnGlnGlyLeuIleThr 330
 DB 1074 TGGGATGTGAGAAAGCTGCAAGATTTATTCATGTTCAGATCAACATGCTCTCTCACA 1133
 QY 331 LeuGlyTyrIleHisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeuHis 350
 DB 1134 CTGGGATGATCCATATACACCCCAAGCAAGGAGATTCCTGTCAGGTGATCTCCAC 1193
 QY 351 ThrHisCysSerTyrGlnMetMetLeuProGluSerValAlaIleValCysSerProLys 370
 DB 1194 ACTCACTGCTCTTCCACAGATATGTGCCAGAGGCATATGCTATGTGTTCCCAAG 1253
 QY 371 PheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeuGluGlnIleSerSerCys 390
 DB 1254 CATTAAGACACCGGATTTTTCAGCTCACCAACGCTGGATCTTGAGGTTCTACTGT 1313
 QY 391 ArgGlnLySgLyPheHisProHisSerLySAspProProLeuPheCysSerSerHis 410
 DB 1314 AAAAGAAAGGCTTCATCTCTTACACCAAGGCCCAAGCTGTATAGTATATGACGCAAT 1373
 QY 411 ValThrValValAspArgAlaValThrIleThrAspLeuArg 424
 DB 1374 GTGTTAGTAAAGACATAAAACAATGCTGTGATCTGAGG 1415
 RESULT 7
 ID ABRK13102 standard; cDNA, 1973 BP.
 AC ABRK13102;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human secretory polynucleotide (spcm) cDNA (001978-1).
 XX
 KW Signal peptide; transmembrane domain; human; spem; ss; gene; 001978.1;
 KW antiatherosclerotic; antiatherosclerotic; antidiabetic;
 KW antiinflammatory; cytoprotic; anti-HIV; antiallergic; antidiabetic;
 KW nephrotropic; antihypert; hepatotropic; neuroprotective;
 KW osteopachic; antihumetic; antiarthritic; dermatological; cancer;
 KW immunosuppressive; anticancer; ophthalmological; vulnery; gout;
 KW anticonvulsant; cerebroprotective; nootropic; antiparkinsonian; virucide;
 KW antibacterial; cell proliferative disorder; arteriosclerosis;
 KW atherosclerosis; psoriasis; immune system disorder; inflammation;

QY 355 TygInMetLeuProGluSerValAlaIleValCysSerProLysPheGlnGluThr 374
Db 1496 TATCAACTCATGTTGCCAGAGGACCATGTCATGTTGCTCACAAAGCATTAAGACACT 1555
QY 375 GlyPhePheLysLeuThrAspHisGlyLeuGlnGluIleSerSerCysArgGlnLysGly 394
Db 1556 GGCATCTTCAGGCTCACCAATGCTGGCATGCTTGGATGCTTGTAAAAAAGGGC 1615
QY 395 PheHisProHisSerLysAspProPhePheCysSerCysSerHisValThrValVal 414
Db 1616 TTTCATCCACACACACAGAGCCAGCTGTTCACTATATGCAACATGCTGTGTA 1675
QY 415 AspArgAlaValThrIleThrAspLeuArg 424
Db 1676 GACATAAAAAATATGCTGTTGATCTGAGC 1705
RESULT 8
AA159567
ID AA159567 standard; cDNA; 4052 BP.
AC AA159567;
XX 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 1770.
DE Human, noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX Homo sapiens.
OS WO200153312-A1.
PN 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
PF 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-0048725.
PR 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-0063450.
PR 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HSE-) HXSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM40411.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Claim 1; SEQ ID NO 1770; 10078bp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA158642-AA162213) with noctropic;
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 4052 BP; 1156 A; 905 C; 910 G; 1081 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 9,1e-101 Length: 4052
Score: 1229.50 Matches: 232
Percent Similarity: 73.97% Conservative: 72
Best Local Similarity: 56.45% Mismatches: 92
Query Match: 55.68% Indels: 15
DB: 4 Gaps: 2
US-09-831-452-1 (1-424) x AA159567 (1-4052)
QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValArgAlaLeuSerGln 20
Db 872 ATGCCCTGACCATACAGATGTTCCCTAAGCCAGAGGAGCTCCGTCCTTAAGCAAG 931
QY 21 LeuGlySerAlaValAlaGluValAsnGluIleProProArgArgTyrPheArgSerGly 40
Db 932 CTGGTTGTAATATACCATCCATGTAAGACATCCAGCTGATCTTGTGCTGGA 991
QY 41 ValGluIleAlaArgMetAlaSerIleTyrSerGlnGluGlyAsnIleGluHisAlaPhe 60
Db 992 GTAGACATGACAGATGCGCTGCTGCTGTTTGGAGAGAGAAATTTGAAATAGCTTT 1051
QY 11eLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
Db 1052 GTTCTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1111
QY 81 LysSerAlaValIleProGluLysLysAspThrValLysLeuLysGluIleAlaPhe 100
Db 1112 CAGCAATGTGACGTACCTGTAAGAGAGATTAATTAAGAGAACTAAGAGAGATTCATTG 1171
QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysLysArgTyrThrLysGluTyrThrGln 120
Db 1172 CCAAGACAGATGATTAAGAAAGACACCTTTAAAGAAATATTAAGCTAAGATCAAGAA 1231
QY 121 TyrAsnGluGluLysLysLysGluAlaGluLysLysAlaArgAsnMetAlaIleGlnGln 140
Db 1232 TATTTGCMAAGCAAAACAAATTAAGCTGAAATTTCCMAAAATTTGAGCATCAGAGA 1291
QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnGlnGlnGln 160
Db 1292 TTGATAGAGCGAAGAAAGAGCGGATGCTCGATGCGCCACACAGCATTAATCGGAG 1351
QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGlnLysGluLysGluLysArgLys 180
Db 1352 CAGTTCTGTTTTCGAGATCACTCAAGAGAGATTAAGCCGAGGTCAATTCGGA 1411
QY 181 IleValGlnGluPheGlyLysValAspProGlyLysGlyGlyProLeuValProAspLeu 200
Db 1412 AGTCAGCAAACTCAGGG-----CTGTCAGAGCAGCAAT 1444
QY 201 GluLysProSerLeuAspValPheProThr-----LeuThrValSer 214
Db 1445 GATGGAGCGCTTGTGCTGCTTTCACACACACAGAACAATTCCTGCTGATATTT 1504
QY 215 SerIleGlnProSerAspCysHisThrThrValAlaGProAlaLysProProValAlaAsp 234
Db 1505 GCAGATCAACCTTAATAAAGTATGCAACCAATTATGTCAGCACTCCCTCTTAAC 1564
QY 235 ArgSerLeuLysProGlyAlaLeuSerAsnSerGlnSerIleProThrIleAspGlyLeu 254
Db 1565 AAGGCGCTTAAGCCAGAGCTGCTCTTAAGTGTGTCAGAAATTAAGTGTGAAGACTG 1624
QY 255 ArgHisValAlaValProGlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaAsn 274


```
|||||
Db AGGCCCTTAACGCCAGCTGCTACTCTACTGCTTCAGATTTAGTGTGAAGGACTG 1624
Qy ArgHisValValValProGlyArgLeuGlyProGlnPheLeuGlnLeuAlaSerAlaAsn 274
Db 1625 CGAGTGTAGTATTTTCCAGGAAGATCTTGGCCAAATTTCTGCACACTGGCAGATCTAAT 1684
Qy 275 ThrAlaArgGlyValGluThrCysGlyIleLeuGlyLeuGlyLeuMetArgAsnGluPhe 294
Db 1685 ACAGTGTAGAGGAATATGAACCTGTGAATCTGTGTGAAAAGTGCACACATTAATGAATTT 1744
Qy 295 ThrIleThrHisValLeuIleProGlyGlnSerIleGlySerAspIleThrCysAsnThrGlu 314
Db 1745 ACTATTAACCATGTAAATTTGTCCAAAGCAGTCTGGGACCAACGATTTGTGACATGAG 1804
Qy 315 AsnGluGluGluLeuPheLeuIleGlnAspGlnGlnIleLeuIleThrLeuGlyTyrIle 334
Db 1805 AATGTAGAGGAATATTTCATGTTTCAAGATCAACATGATCTCCACACTTAGAGTGGATC 1864
Qy 335 HisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeuHisThrHisCysSer 354
Db 1865 CATACACATCCCATCAAACTGATTTTATCCAGCGTTGATCTTCACTCACTGTTCC 1924
Qy 355 TyrGlnMetMetLeuProGluSerValAlaIleValCysSerProGlyPheGlnGluThr 374
Db 1925 TATCACTCATGTTGTCAGAGGAGCCATTGCTTGTCTCACCAGCATTAAGACACT 1984
Qy 375 GlyPhePheLeuThrHisPheHisGlyLeuGluGluIleSerSerCysArgGlnGly 394
Db 1985 GGCACTCTTCCAGGCTCACCAATGCTGGCATGCTTGAAGTTCCTGTATAAAAAAGGCG 2044
Qy 395 PheHisProHisSerIleCysAspProProLeuPhe 405
Db 2045 TTTCATCCACACACCAAGAGGCCACAGCGTGTTC 2077

RESULT 10
AAS85218
ID AAS85218 standard; cDNA; 1236 BP.
XX
AC AAS85218;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #21022.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX OS
XX WC200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WC-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WP1: 2001-639362/73.
XX
XX P-PSDB; ABG21031.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX
XX Claim 1; SEQ ID NO 21022; 103pp; English.
XX
```

```
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp://ipo.int/pub/published_pct_sequences
XX
SQ Sequence 1236 BP; 331 A; 317 C; 338 G; 250 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.52e-50 Length: 1236
Score: 1102.50 Matches: 219
Percent Similarity: 91.34% Conservative: 13
Best Local Similarity: 86.22% Mismatches: 17
Query Match: 49.93% Indels: 5
DB: Gaps: 3
US-09-831-452-1 (1-424) x AAS85218 (1-1236)
Qy 102 LysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyr----- 118
Db 172 GAGGCTCAGAGCAAGAGAGGATCTTAATCTCAATACCCCATTAACACGCGCATCATC 231
Qy 119 ThrGluTyrAsnGlu---GluLysLysGlyGluAlaGluGluLeuAlaArgAsnMetAla 137
Db 232 ACCAACTGGATGACCTGGAGAAAGAAAGAAAGAGAGAAATGGCCCGGAACATGGGC 291
Qy 138 IleGlnGlnGluLeuGluLysGlnArgValAlaGlnGlnLysGlnGlnLeu 157
Db 292 ATCCAGCAGACACTGGAAAAAGAAAAACAGAGGTAGACACACAGACAGCAGCATGG 351
Qy 158 GluGlnGlnPheHisAlaPheGluGluMetIleArgAsnGlnGluGluLysGlu 177
Db 352 GAACAGGAACAGTTCATGCTTCGAGAGATGATCCGAAACAGAGACTAGAAAAAGG 411
Qy 178 ArgLeuLysIleValGlnGlnPheGlyLysValAspProGlyLeuGlyGlyProLeuVal 197
Db 412 CCACTGAATAATGTACAGAGAGTTTGGGAAGTACACCTCGGCTAGTGGCCGCTAGG 471
Qy 198 ProAspLeuGluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGln 217
Db 472 CCGAATTTGAGAAAGCCCTCTTAATGTGTCCACCTTAACAGTCTCATCCATACAG 531
Qy 218 ProSerAspCysHisThrThrValArgProAlaLysProProValValAspArgSerLeu 237
Db 532 CCTTCAGACTGTACCAACTGAAGCCAGCTAAGCCACTGTGTGTGACAGGTCCTTG 591
Qy 238 LysProGlyAlaLeuSerAsnSerGlnSerIleProThrIleAspGlyLeuArgHisVal 257
Db 592 AAACCTGGAGCACTGAGCAACTCAGAAAGTATCCACATGATGATGGCCATGG 651
Qy 258 ValValProGlyArgLeuGlyProGlnPheLeuGlnLeuAlaSerAlaAsnThrAlaArg 277
Db 652 GGGGTGCGGGGGGGGTGGCCACAGATTCTCCAGTTAGCCAGTGCACACCTGCCG 711
Qy 278 GlyValGluThrCysGlyIleLeuGlyLysLeuMetArgAsnGluPheThrIleThr 297
Db 712 GGAGTGAACATGTGGAAATTTCTGTGTGAAAAAGTATGAGAAATTAATTCATTAAC 771
```

Qy	298	HisValIleuIlePProLYSGInSerAlaGlySerAspTYrCYAsnThrGluAsnGluGlu	317
Db	772	CATGTTCTCATCCCAAGCAAGAAAGTCGGGTCTGATTACTGCAGACAGAGACGAGAA	833
Qy	318	GluLeuPheIleuIleGlnAspGInGInGlyIleuIleThrLeuGlyTYrIleHisThrHis	337
Db	832	GAACCTTTCCTCATACAGAGATCAGCGGGCTCATACACTGCAGTTCATGTAAAGC	891
Qy	338	ProThrgIn--ThraIaPheIaSerSerValaIaPheHis	350
Db	892	AATTCGTAGCTCTCCGAGACGATGAGAGTCCGCTGGAATTCAT	933
	RESULT 11		
	AA161352/c		
	ID	AA161352 standard; cDNA; 1318 BP.	
XX	AA161352;		
XX	AC		
XX	DT	22-OCT-2001 (first entry)	
XX	DE		
XX	Human	polynucleotide SEQ ID NO 5341.	
XX	Human;	nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral	nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's;	Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic	lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	leukemia;	thrombolytic; drug screening; arthritis; inflammation;	
XX	chemokine;	ss.	
XX	Homo sapiens.		
XX	WO200153312-A1.		
XX	26-JUL-2001.		
XX	26-DEC-2000;	2000WO-US034263.	
XX	23-DEC-1999;	99US-00471825.	
PR	21-JAN-2000;	2000US-004818725.	
PR	25-APR-2000;	2000US-00552317.	
PR	20-JUN-2000;	2000US-00598042.	
PR	19-JUL-2000;	2000US-00620312.	
PR	03-AUG-2000;	2000US-00653450.	
PR	14-SEP-2000;	2000US-00662191.	
PR	19-OCT-2000;	2000US-00693036.	
PR	29-NOV-2000;	2000US-00727344.	
XX	(HSE-) HYSEQ INC.		
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,		
PI	Wang Y, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA,		
PI	Zhou P, Goodrich R, Drmanac RT;		
XX	WPI: 2001-442253/47.		
DR	P-PSDB; AAM42196.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders such		
PT	as central nervous system injuries.		
XX	Claim 1; SEQ ID NO 5341; 10078bp; English.		
XX	The invention relates to human nucleic acids (AA157798-AA161369) and the		
CC	encoded polypeptides (AAM38642-AA42213) with nocotropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

Sequence 1318 BP; 354 A; 275 C; 257 G; 432 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	1,73e-73
Score:	90.50
Percent Similarity:	68.63%
Best Local Similarity:	51.25%
Query Match:	41.65%
OS:	4
Gaps:	2

US-09-831-452-1 (1-424) X AAT61352 (1-1318)

77 H1SARGASPTYRLLYSERAlAValIlePROGLULYSLSASPTHRValLLYSLYLEuLYS 96

Db 1316 CACGAGATTACCAGCAATGTCAGTACTGAAAGCAGGATATTATGAGAACTGAAG 1257

97 GluLeuAlaPheProLysAlaGluGluLeuLysAlaGluLeuLeuLysArgTyrThrLys 116

Db 1256 GAGATTGCATTCCTCCAGGACAGATGAATTGAAAAACGACCTTTTAAAGAAATATAACGTA 11

117 GluTyrThrGluTyrAsnGluGluLysLysGluAlaGluGluLeuAlaArgAsnMet 13

Db 1196 GAATACCAAGATATTTCGAAGCAAAACAATATATAAGCTGAAATTCTCAAAAAATTG 1137

QY 137 AlaIleGlnGlnIuLeuGluLysGlnArgValAlaGlnGlnLysGlnGln 156

Db 1136 GAGCATCAGAGATTGATAGAGGCAGAAAGGAAGCGATTGCTCAGATGCCGCCAGCAGCAG 107

157 LeuGIuGIuGIuPheHISAlaPheGIuGIuMetIleArgASnGIuGIuLeuGIuLys 176

Db 1076 CTAGATCGAGCAGTTTCTGTTTTTCGAGATCACTCAAGACCAAGAGTTAGCCCCGA 101

177 GluArgLeuLysIleValGlnGluPheGlyLysValAspProGlyLeuGlyGlyProLeu 196

Db 1016 GGTCAATGCGAAGTCAGCAAACCTCAGG-----CTG 984

QY 197 ValProaspLeuGlulysProserLeuaspValpneProThr----- 210

Db 983 TCAGAGCAGATTGATGGGAGCACTTGTCTCTGCTTTCCACACACCAGAACATTCCTTG 924

```

211 LeUlnrvalserseIllegInProsetAsprCybnIsImIInValmGrIoMaLayStC 230
      :::::          :::::

```

Db 923 CTGAATGTAATTGCAGATCAACCTATAATAAGTGAAGGCACCAATATATGTAAGCCATCTT 884

231 Pro-ValValAspArgSerLeuLysPheGlyAlaLeuSerAsnSerGluSerIlePheCm 230

[illegible][illegible][illegible]

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

[illegible][illegible][illegible]

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

504

QY 350 sThHisCySerTyArgInMeMetLeuProGluSerValAlaIleValCysSerProLy 370
 Db 503 CACCACTGTTCCATCACTCATGTTGCCAGAGCCATTGCCATTGTTGCTCACCAA 444
 QY 370 sPhGlnGluThrGlyPhePheIysLeuThrAspHisGlyLeuGluGluIleSerSerCy 390
 Db 443 GCATTAAGACACTGCGCATCTTCACGCTCCACCAATGCGCATGCTTGTGAGTTCTGCTTG 384
 QY 390 sArgGluIysGlyPheHisProHisSerTyAspProGluPhePheCysSer 410
 Db 383 TAAAAAAGAGGCTTTTCATCCACACACCAAGAGCCAGGCTGTCAGTATATGCAAA 324
 QY 410 sValThrValAlaAspArg-AlaValThrIleThrAspLeuArg 424
 Db 323 ATGTGTTGTAAGACATATAATATGTTGTGATCTGAGG 279

RESULT 12
 ID ABL03899 standard; cDNA; 1611 BP.
 AC ABL03899;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 6179.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 XX MO200171042-42.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001.655860/75.
 DR P-PSDB; ABB59796.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 6179; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB163051), expressed DNA
 CC sequences (AB01840-AB16175) and the encoded proteins (ABB57737-
 CC ABB572022). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1611 BP, 428 A; 439 C; 424 G; 320 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.13e-69 Length: 1611
 Score: 871.50 Matches: 188
 Percent Similarity: 62.10% Conservative: 84
 Best Local Similarity: 42.92% Mismatches: 119
 Query Match: 39.47% Indels: 47
 DB: 4 Gaps: 13

US-09-831-452-1 (1-424) x ABL03899 (1-1611)
 QY 5 GlyAspValSerLeuProProGluAspArgValAlaArgAlaLeuSerGlnLeuGlySerAla 24
 Db 379 GGGAGCGTGAG-CCGAGAGAGAGATGAGACCTTCGATCGCGCAACTCG 432
 QY 25 ValGluValAsnGluAspIleProProArgArgTyPheArgSerGlyValGluIle 44
 Db 433 ATCGAGGTGAGCAAGACATGCCAGTCAAGCGGTCTACCGCTTGCGGACCGAGATGCTG 492
 QY 45 ArgMetAlaSerIleTySerGluGluValIleGluHisAlaPheIleLeuTyra 64
 Db 493 CGCATGCCCAATGTTTACCTCGCGGAGGCAACACGAGAACGATTCATCTGTAATTG 552
 QY 65 LysTyrlleThrLeuPheIleGluIysLeuProLysHisArgAspTyrlleSerAlaVal 84
 Db 553 CGCTACATACCTCTTCAATCAGAAAGATCCGCGACACCCGATTTAGCGAC---GTC 609
 QY 85 IleProGluLysLeuAspThrValIysIysLeuLysGluIleAlaPheProLysAlaGlu 104
 Db 610 AAGGCCAGGTGCGGATATTAACAGAAAGATTAAAGACAGATTAATGCGCCACGACGAG 669
 QY 105 GluLeuLysAlaGluLeuLeuLysArgTyrlleLysGluTyrlleThrGluTyrlleAsnGlu 124
 Db 670 AAACGTGAGCCAGACTACTTAACGCACTACAAAGGAGATGAGAGATTCTGCGCTAGC 729
 QY 125 LysIysLysGluAlaGluLeuLeuAlaArgAsnMetAlaIleGlnGlnIleLeuGluLys 144
 Db 730 AAGGAGCGCGAA-----CGAGTCAAGAACTGAACGCGGACGCGG 774
 QY 145 GluLysGlnArgValAlaGlnGlnLysGlnGlnIleGlnGluGlnIlePheHisAla 164
 Db 775 GAAGAGGAACGT-----CAGCGCAAAAGGAAGGAGAGACAGGTTCAAGTGT 825
 QY 165 PheGluMetIleArgAsnGlnIleGluLysGluLysGluLysIleValGlnGlu 184
 Db 826 ATCCCATCCCTCATC-----CGGCGCAATTTGCACGTGCTC----- 861
 QY 185 PheGlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeuGluLysProSer 204
 Db 862 -----ATCGACGAAGGC---AACCAACCATCGGCTCCGACCTGGCTGCTTAC 909
 QY 205 LeuAspValPheProthr----- 210
 Db 910 CAGGTGTTTACCCCAATGACTTTCCACGCGCGCCCATGCGACCTGCCGGATCGGGA 969
 QY 211 -----LeuThrValSerSerIleGlnProSerAspCysHisThrThrValArgProAla 228
 Db 970 CTGCTGCTCGCGGCTCCAGCAAGAGAGCGCGCAACAAACCAATCCCAAGCAACG 1029
 QY 229 LysProProValValAspArgSerLeuLysProGlyAlaLeuSerAsnSerIle 248
 Db 1030 -----TTCATGCAATCAGAGCCA---TCTTACACCGCAACATCTCTG 1074
 QY 249 ProThrIleAspGlyLeuArgHisValValValProGlyArgLeuCysProGlnPheLeu 268
 Db 1075 ---CTGGCGGCTCTTGGCGCTCGTTTACGTCGCGCGGAGACCATGAGAGTGTTCG 1131
 QY 269 GlnLeuAlaSerAlaAsnThrAlaArgGlyValGluThrCysGlyTyrlleuCysGlyLys 288
 Db 1132 AAGCTGCGCTTGGCCCAACCTCGAAACATCGAGACATGCGGCTACTGCTGCTCAT 1191
 QY 289 LeuMetArgAsnGluPheThrIleThrHisValIleIleProLysGlnSerAlaGlySer 308
 Db 1192 CTGTCCCAAGACAGCTGATCATCTCATCATCATTTAGCCCGGACAGCAAGAACTCGG 1251
 QY 309 AspTyrlleCysAsnThrGluAsnGluGluLeuPheLeuIleGlnAspGlnGlnIleu 328
 Db 1252 GACAGCTGCAACACGATGACAGAGGAGAGATTCGATGTCAGGACCAAGTGCAGCTG 1311
 QY 329 IleThrLeuGlyTyrlleHisThrHisProThrGlnThrAlaPheLeuSerSerValAsp 348


```

Db      |||||
1383  ACTGCATTTTATCCAGCGTGTGATCTTCACACTCAGTTCCCTATCCATCATGTTCCCA 1442
Qy      |||||
361  GluSerValAlaIleValCysSerProLysPheGlnIuPhrgIyPhePheLysLeuThr 380
Db      |||||
1443  GAGGCCATTCGCATTTGTTGTCACCAAGCATAAAGACACTGGCATCTTCAGGCTTACC 1502
Qy      |||||
381  AspiHisGlyLeuGluGluIleSerSerCysArgGlnIySglYpHeH:SPROH:SerLys 400
Db      |||||
1503  AATGCTGACATGCTTGAGGTTTCTGCTGTGAAAAAAGGCTTTTCATCCACACACCAAG 1562
Qy      |||||
401  AspProPheLeuPheCysSerCysSerHisValThrValValAspaGalaValThrIle 420
Db      |||||
1563  GAGCCCAAGGCTGTTCAGTATATGCAACATGTGTGGTAAAAAGCATMAAAATTAATTGTG 1622
Qy      |||||
421  ThrAspLeuArg 424
Db      |||||
1623  TTGGATCTGAGG 1634

RESULT 14
AAA69708
ID      AAA69708 standard; cDNA; 1041 BP.
XX      AAA69708;
AC      AAA69708;
XX      07-NOV-2000 (first entry)
DT
XX      Human ovarian carcinoma antigen polynucleotide SEQ ID NO:18.
DE
XX      Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;
KM      tumour antigen; identification; cytostatic; gene therapy; vaccine; ss.
XX      Homo sapiens.
OS
XX      WO200036107-A2.
PN
XX      22-JUN-2000.
PD
XX      17-DEC-1999; 99WO-US030270.
PF
XX      17-DEC-1998; 98US-00215681.
PR      17-DEC-1998; 98US-00216003.
PR      23-JUN-1999; 99US-00338933.
PR      24-SEP-1999; 99US-00404879.
XX
XX      (CORI-) CORIXA CORP.
PA
XX      Mitcham JL, King GE, Algate PA, Frudakis TN;
PI      WPI; 2000-431589/37.
PN
XX
XX      Immunogenic portion of an ovarian carcinoma protein and the nucleic acid
PT      encoding it, useful for the diagnosis, prevention and treatment of
PT      cancer, preferably ovarian cancer.
XX
XX      Claim 1; Fig 1; 299pp; English.
PS
XX      The present invention describes an isolated polypeptide comprising an
CC      immunogenic portion of an ovarian carcinoma protein (or its variants).
CC      Ovarian carcinoma proteins, and polynucleotides encoding them, have
CC      cytostatic activity and can be used in gene therapy and vaccines. Ovarian
CC      carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful
CC      for the prevention, diagnosis and treatment of cancer, preferably ovarian
CC      cancer. AAA69691 to AAA70077 and AAB12552 to AAB12557 represent human
CC      ovarian carcinoma polynucleotides and proteins used in the
CC      exemplification of the present invention
XX
SQ      Sequence 1041 BP; 312 A; 223 C; 198 G; 307 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 6,21e-59 Length: 1041
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0
DB: 3 Gaps: 0
US-09-831-452-1 (1-424) x AAA69708 (1-1041)
Qy      |||||
285  LeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValIleuLeuProLysGln 304
Db      |||||
1  CTCTGTGCAAACTGATGAGAAATGATTAATTCACATTACCCATGTTCTCATCCCAAGCA 60
Qy      |||||
305  SerAlaGlySerAspTyrCysAsnThrGluAsnGluGluIuPhePheLeuIleGlnAsp 324
Db      |||||
61  AGTGTGAGTCTGATTAATTCGACACACAGAACAGAAAGAACTTTCTCTTACAGGAT 120
Qy      |||||
325  GlnGlnGlyLeuIleThrLeuGlyTyrPleHisThrHisProThrGlnThrAlaPheLeu 344
Db      |||||
121  CAGCAGGCGCTCATGACACTGGGCTGGATTCATCTCACCCACACAGCCGCTTCTC 180
Qy      |||||
345  SerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuProGluSerValAla 364
Db      |||||
181  TCCAGTGTGACCTACACACTCAGCTCTTACAGATGATTTGCCAGAGTCAGTAGCC 240
Qy      |||||
365  IleValCysSerProLysPheGlnIuPhrgIyPhePheLysLeuThrAspiHisGlyLeu 384
Db      |||||
241  ATTGTTTGCTCCGCCAAGTTCAGGAAGACTGATTTAAACTAATGACATGACATGACTA 300
Qy      |||||
385  GluGluIleSerSerCysArgGlnIySglYpHeH:SPROH:SerLysAspProLeu 404
Db      |||||
301  GAGGAGATTTCTTCTGTGCGCAGAAAGATTTTCATCCACAGCAAGATCCACTCTG 360
Qy      |||||
405  PheCysSerCysSerHisValThrValValAspaGalaValThrIleThrAspLeuArg 424
Db      |||||
361  TTCTGTAGCTGCAGCCAGCTGCTGTGTGGACAGAGATGACATCATCAGACCTTCCA 420

RESULT 15
ABN72602
ID      ABN72602 standard; DNA; 1041 BP.
XX
XX      ABN72602;
AC
XX
XX      02-JUL-2002 (first entry)
DT
XX
XX      Ovarian carcinoma antigen polynucleotide #18.
DE
XX      Human; immunostimulant; cytostatic; cancer; ovarian carcinoma; ds.
KM
XX      Homo sapiens.
OS
XX      WO200206317-A2.
PN
XX
XX      24-JAN-2002.
PD
XX      17-JUL-2001; 2001WO-US022635.
PF
XX      17-JUL-2000; 2000US-00617747.
PR      10-AUG-2000; 2000US-00636801.
PR      20-SEP-2000; 2000US-00667857.
PR      04-APR-2001; 2001US-00837271.
PR      18-JUN-2001; 2001US-00884441.
XX
XX      (CORI-) CORIXA CORP.
PA
XX      Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR,
PI      Reed SG, Vedrick TS, Carter D, Hill P, Albore E;
XX      WPI; 2002-164781/21.
DR
XX
XX      Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT      protein or its variants, useful for stimulating an immune response in a
PT      patient and treating ovarian cancer.
XX
XX      Example 1; Page 98; 408pp; English.

```

CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents DNA related
CC to the invention

XX Sequence 1041 BP; 312 A; 223 C; 198 G; 307 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 6,21e-59 Length: 1041
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0
DB: 6 Gaps: 0

US-09-831-452-1 (1-424) x ABN72602 (1-1041)

QY 285 LeuCyseGlyLysLeuMetArgAsnGluPheThrIleThrHisValIleuIleProLysGln 304
DB 1 CTGTGTGAAAGCTGATGAGGAATGATTAACATTAACCATGTCTCATCCCAAGCAA 60
QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPheLeuIleGlnAsp 324
DB 61 AGTGTCTGGTCTGATTACTGCAACACAGAGAGAGAACTTTCTCATACAGAT 120
QY 325 GlnGlnGlyLeuIleThrLeuGlyTyrPleHisThrHisProThrGlnThrAlaPheLeu 344
DB 121 CAGCAGGGCCTCATCACACTGGGCTGATTCATCTACCCCAACAGACCGGCTTCTC 180
QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetLeuPProGluSerValAla 364
DB 181 TCCAGTGTGACCTACACACTCACTGCTTACCAGATGATGTCCAGAGTCACTAGCC 240
QY 365 IleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeu 384
DB 241 ATTGTGTCTCTCCCAAGTTCAGAAATCGAATCTTTAACTAACTGACCATGACCTA 300
QY 385 GluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProLeu 404
DB 301 GAGGAGATTTCTTCTCTGCGCCAGAAAGATTTCATCCACAGCAAGATCCACTCTG 360
QY 405 PheCysSerCysSerHisValThrValValAspArgAlaValThrIleThrAspLeuArg 424
DB 361 TTCTGTAGTGCAGCCACGTGACTGTGTGACAGACAGTGAACCATCACAGACTTCGA 420

Search completed: August 23, 2004, 19:20:59
Job time : 559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 23, 2004, 18:57:11 ; Search time 102 Seconds

(without alignments)
2306.855 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208

Sequence: 1 MSBHDVSLPDEPRVALSQ.....FCGSHVTVDRAVTTIDLR 424

Scoring table:

BIOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgm2_1/USPTO.spool/US09831452/runat_17082004_151200_23681/app_query.fasta_1.583
-DB=Issued_Patents_NA -QMT=fasta -SUF=TX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.ccd1
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09831452 @CGN 1.1 69 @runat_17082004_151200_23681 -NCPU=6 -ICPU=3
-NO MMAP -LARGOJURY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA.*
1: /cgm2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgm2_6/ptodata/2/ina/5A COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	2003	4	US-09-665-479A-1
2	758	34.3	1041	4	US-09-404-879A-18
3	758	34.3	1041	4	US-09-358-933-18
4	758	34.3	1041	4	US-09-215-681-18
5	758	34.3	1041	4	US-09-216-003A-19
6	758	34.3	1043	4	US-09-404-879A-18
7	758	34.3	1043	4	US-09-338-933-19
8	758	34.3	1043	4	US-09-215-681-19
9	758	34.3	1043	4	US-09-216-003A-19
10	131.5	6.0	4266	4	US-09-651-011A-3
11	127.5	5.8	3095	6	5231168-1
12	125	5.7	6856	4	US-09-566-921-42

13	122.5	5.5	3798	4	US-09-688-186B-9	Sequence 9, Appl1
14	122.5	5.5	3798	4	US-09-291-417D-9	Sequence 9, Appl1
15	118	5.3	3972	4	US-09-645-456A-6	Sequence 6, Appl1
16	118	5.3	3972	4	US-09-425-324A-6	Sequence 6, Appl1
17	118	5.3	3972	4	US-09-645-791-6	Sequence 6, Appl1
18	118	5.3	3996	4	US-09-645-456A-2	Sequence 2, Appl1
19	118	5.3	3996	4	US-09-425-324A-2	Sequence 2, Appl1
20	118	5.3	3996	4	US-09-645-791-2	Sequence 2, Appl1
21	118	5.3	4055	4	US-09-688-186B-10	Sequence 10, Appl1
22	118	5.3	4055	4	US-09-291-417D-10	Sequence 10, Appl1
23	117	5.3	2133	4	US-09-107-532A-1413	Sequence 1413, Ap
24	116.5	5.3	5934	4	US-09-418-710-2	Sequence 2, Appl1
25	114.5	5.2	7812	3	US-09-168-590-1	Sequence 1, Appl1
26	114	5.2	2457	4	US-09-134-000C-2950	Sequence 2950, Ap
27	114	5.2	8146	4	US-09-576-594-725	Sequence 725, App
28	114	5.2	11917	4	US-09-566-921-32	Sequence 32, Appl
29	113.5	5.1	2603	4	US-09-620-312D-903	Sequence 903, App
30	113.5	5.1	1664976	4	US-08-916-421B-1	Sequence 1, Appl1
31	113	5.1	3974	3	US-08-467-504-3	Sequence 3, Appl1
32	113	5.1	11236	1	US-07-853-913-1	Sequence 1, Appl1
33	112.5	5.1	2896	4	US-09-266-225D-9	Sequence 9, Appl1
34	112	5.1	10136	1	US-08-353-700-2	Sequence 2, Appl1
35	112	5.1	10136	5	PCT-US95-16216-2	Sequence 2, Appl1
36	111.5	5.0	1131	6	5180810-3	Patent No. 5180810
37	111.5	5.0	1784	6	5180810-2	Patent No. 5180810
38	111.5	5.0	4722	4	US-08-979-608A-14	Sequence 14, Appl
39	111.5	5.0	4722	4	US-09-517-849-14	Sequence 14, Appl
40	111.5	5.0	4722	4	US-09-616-289-14	Sequence 2819, Ap
41	111	5.0	714	4	US-09-134-001C-2819	Sequence 2264, Ap
42	111	5.0	3543	4	US-09-543-681A-2264	Sequence 209, App
43	111	5.0	51259	3	US-08-781-891-209	Sequence 209, App
44	111	5.0	51259	4	US-09-618-166-209	Sequence 1, Appl1
45	110.5	5.0	3664	4	US-09-555-554-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-665-479A-1
; Sequence 1, Application US/09665479A
; Patent No. 6673570
; GENERAL INFORMATION:
; APPLICANT: Itoh, Fumiko
; APPLICANT: Itoh, Susumu
; APPLICANT: Haldin, Carl-Henrik
; APPLICANT: Ken-Di, Peter
; TITLE OF INVENTION: SMAD ASSOCIATING POLYPEPTIDES
; FILE REFERENCE: L00461.70096.US
; CURRENT APPLICATION NUMBER: US/09/665,479A
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US 60/154, 846
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)..(1334)
US-09-665-479A-1
Alignment Scores:
Pred. No.: 4,49e-243 Length: 2003
Score: 2208.00 Matches: 424
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-831-452-1 (1-424) X US-09-665-479A-1 (1-2003)

QY 1 MetSerAspHisGlyAspValSerLeuProGluAspArgValAlaLeuSerGln 20
 Db 63 ATGTCTGACCATGAGATGTGAGCTCCGCCCGGAGACCGGCTGAGGCTCTCTCCAG 122
 QY 21 LeuGlySerAlaValGluValAsnGluAspIleProProArgArgTyrThrArgSerGly 40
 Db 123 CTGGGTAGTGGGTAGGTGATGAAAGCATTCACCCCGTGGTACTCGCTCTGGA 182
 QY 41 ValGluIleIleArgMetAlaSerIleTyrSerGluGluGlyAsnIleGluHisAlaPhe 60
 Db 183 GTTGAAGATTATCCGATGCGATTCATTACTTGAAGAGGACACTTGAACATGCTTC 242
 QY 61 IleLeuTyrAsnLysTyrIleThrLeuPheIleGluLysLeuProLysHisArgAspTyr 80
 Db 243 ATCTCTATACAAAGATATATCAGCTCTTATTGAAAACATACCAAAAACATGAGATTAC 302
 QY 81 LysSerLeuAlaValIleProGluLysLysAspThrValLysLysLeuGluIleAlaPhe 100
 Db 303 AAATCTGCTGTCTATCTCTGAAAAGAAACACACGTAAGAAATTAAGAGATTCATT 362
 QY 101 ProLysAlaGluGluLeuLysAlaGluLeuLysArgTyrThrLysGluTyrThrGlu 120
 Db 363 CCCAAAGCAGAGAGCTGAGAGCCAGAGCTGTTAAACGATATACCAAGAAATACAGAA 422
 QY 121 TyrAsnGluGluLysLysGluAlaGluGluLeuAlaArgMetAlaIleGlnGln 140
 Db 423 TATATTAAGAAAAGAAAGAAAGAAAGCAAGAAATGCGCCGGAACATGGCATCGACA 482
 QY 141 GluLeuGluLysGluLysGlnArgValAlaGlnGlnLysGlnGlnLeuGlnGln 160
 Db 483 GAGCTGAAAAAGAAAAACAGAGGTAGCAACAAGAGCAGCAGCAATGGAAACAGAA 542
 QY 161 GlnPheHisAlaPheGluGluMetIleArgAsnGlnGluLeuGluLysGluAlaGluLys 180
 Db 543 CAGTTCCATGCTCTCGAGAGATGATCCGAAACAGAGCTTAAAGAAAGCCACGAA 602
 QY 181 IleValGlnGluPheGlyLysValAspProGluLysGlyProLeuValProAspLeu 200
 Db 603 ATTGTACAGAGATTGGAAAGTGAAGTACCTGCGCTAGAGTGGCCCGCTAGGCTGACTTG 662
 QY 201 GluLysProSerLeuAspValPheProThrLeuThrValSerSerIleGlnProSerAsp 220
 Db 663 GAGAAAGCCCTCCTAGATGTGTCCCACTTAAAGTCTCAATCCATACAGCCCTTAGAG 722
 QY 221 CysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLysProGly 240
 Db 723 TGTACACAACTGTAGGCCAGCTAGACCTGTGTGAGACAGTCTTGAACCTGGA 782
 QY 241 AlaLeuSerAsnSerGluSerIleProThrIleAspGlyLeuArgHisValAlaPro 260
 Db 783 GCACTGAGCAACTCAGAAAAGTATCCCAATCGATGATTCGCGCATGTGTGTGCT 842
 QY 261 GlyArgLeuCysProGlnPheLeuGlnLeuAlaSerAlaSerThrAlaArgGlyValGlu 280
 Db 843 GGGGGGCTGTGCCACAGTTCCTCAGTTAGCCAGTCCAACTGCCCGGGGAAGTGA 902
 QY 281 ThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeu 300
 Db 903 ACACTGTGAATCTCTGTGAAATCTGATGAGAAATGATTTACCATTAACCATGTCTTC 962
 QY 301 IleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPhe 320
 Db 963 ATCCCAAGCAAAAGTGTGGTCTGATTACTGCAACACAGAGAAAGAAAGAACTTTTC 1022
 QY 321 LeuIleGlnAspGlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGln 340
 Db 1023 CTCTATACGAGATCAGCGAGGCTCTATACACTGGGCTGGATTCTATATCCACACAG 1082
 QY 341 ThrAlaPheLeuSerSerValAspLeuHisThrHisCysSerTyrGlnMetValLeuPro 360
 Db 1083 ACCGGTATTCTCTCAAGTGTGACCTACACACTGCTTACACAGAGATGTTGCCA 1142
 QY 361 GluSerValAlaIleValCysSerProLysPheHisGlnThrGlyPhePheLysLeuThr 380

Db 1143 GAGTCAGAGCATGTTGTTGCTCCCAAGTCCAGAAATGGATTCCTTAACTACT 1202
 QY 381 AspHisGlyLeuGluGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
 Db 1203 GACCATGACACTAGAGGAGATTCTCTCTGTGGCAGAAAGATTTATCCACACAGAG 1262
 QY 401 AspProProLeuPheCysSerCysSerHisValThrValAlaAspAlaValThrIle 420
 Db 1263 GATCACCCTCTGTCTGTGTAGCTGACGACGATGACTGTTGTGACAGAGATGACCATC 1322
 QY 421 ThrAspLeuArg 424
 Db 1323 ACAGACCTTGA 1334
 RESULT 2
 US-09-404-879A-18
 ; Sequence 18, Application US/09404879A
 ; Patent No. 6468546
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Algate, Paul A.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.462C2
 ; CURRENT FILING DATE: 1999-03-24
 ; NUMBER OF SEQ ID NOS: 393
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 18
 ; LENGTH: 1041
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1041)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-404-879A-18
 Alignment Scores:
 Pred. No.: 3,69e-77 Length: 1041
 Score: 758.00 Matches: 140
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 34.33% Indels: 0
 DB: 4 Gaps: 0
 US-09-831-452-1 (1-424) x US-09-404-879A-18 (1-1041)
 QY 285 LeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHisValLeuIleProLysGln 304
 Db 1 CTCTGTGAAAACTGATGAGAAATGAATTTACCATTTACCCATGTCTTCATCCCAACCA 60
 QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGluLeuPheLeuIleGlnAsp 324
 Db 61 AGTGTGGGTCTGATTCTGACACACAGAAACGAAAGAACTTTCTCTATACAGAT 120
 QY 325 GlnGlnGlyLeuIleThrLeuGlyTyrIleHisThrHisProThrGlnThrAlaPheLeu 344
 Db 121 CAGCAGGGCTCTATACACTGGGCTGATTCATCTACCCCAACAGACCGCTTTCTC 180
 QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetValLeuProGluSerValAla 364
 Db 181 TCCAGTGTGACCTTACACACTGCTCTTACCAAGATGATGTGCGAGACTGAGTACC 240
 QY 365 IleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeu 384
 Db 241 ATTGTGCTCTCCCAAGTTCAGAAATCTGATTTCTTAACCTAAGACATGAGACTA 300
 QY 385 GlnGluIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProProLeu 404
 Db 301 GAGGAGATTCTTCTCTGTGGCCAGAAAGATTTATCCACACAGCAAGATCCACTCTG 360

```

QY      405 PheCySerCySeSerHisValThrValAlaSPaRgaIaValThrIleThraSPLeuArg 424
Db      361 TTCGTGAGTGCAGCCACGTAAGTGTGTGGACAGACGAGTGCATCAAGACCTTCGA 420

RESULT 3
US-09-338-933-18
; Sequence 18, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338, 933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1041)
; OTHER INFORMATION: n = A,T,C or G
US-09-338-933-18

Alignment Scores:
Pred. No.:      3.69e-77      Length:      1041
Score:          758.00      Matches:      140
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    34.33%      Indels:      0
DB:             Gaps:      0

US-09-831-452-1 (1-424) x US-09-338-933-18 (1-1041)
QY      285 LeuCySGlyLysLeuMetArgAenGluPheThrIleThrHisValLeuIleProLysGln 304
Db      1 CTCGTGGAAGAAACGATGAGGAATGAATTACCTTCCATGCTTCATCCCAAGCA 60
QY      305 SerAlaGlySerAspTyrCySasnThrGluAsnGluGluLeuPheLeuIleGlnAsp 324
Db      61 AGTCTGGGTCTGATTACTGCAACACAGAACGAAAGAACTTTCTCATACAGAT 120
QY      325 GlnGlnGlyLeuIleThrLeuGlyTTrpIleHisThrHisProThrGlnThrAlaPheLeu 344
Db      121 CAGCAGGGCCTCATCACTGGGCTGGATTCTATCTACCCCAACAGACCGGCTTCTC 180
QY      345 SerSerValAspLeuHisThrHisCySserTyrGlnMetLeuProGlnSerValAla 364
Db      181 TCCAGTGTGCAGCTACACACTGCTCTTACCAAGATGATGTGCCAGAGTCAGTAGCC 240
QY      365 IleValCySerProLysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeu 384
Db      241 ATTGTTGCTCCCAAGTTCAGGAATCGATCTTTAACTAATGACATGACATGAGACTA 300
QY      385 GlnGluIleSerSerCySArgGlnLysGlyPheHisProHisSerLysAspPropoleu 404
Db      301 GAGGAGATTCTCTCTGCTCCCGAAGAAAGATTTCATCCACAGCAAGATCCACTCTG 360
QY      405 PheCySerCySeSerHisValThrValAlaSPaRgaIaValThrIleThraSPLeuArg 424
Db      361 TTCGTGAGTGCAGCCACGTAAGTGTGTGGACAGACGAGTGCATCAAGACCTTCGA 420

RESULT 4
US-09-215-681-18
; Sequence 18, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.

```

```

; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215, 681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1041)
; OTHER INFORMATION: n = A,T,C or G
US-09-215-681-18

Alignment Scores:
Pred. No.:      3.69e-77      Length:      1041
Score:          758.00      Matches:      140
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    34.33%      Indels:      0
DB:             Gaps:      0

US-09-831-452-1 (1-424) x US-09-215-681-18 (1-1041)
QY      285 LeuCySGlyLysLeuMetArgAenGluPheThrIleThrHisValLeuIleProLysGln 304
Db      1 CTCGTGGAAGAAACGATGAGGAATGAATTACCTTCCATGCTTCATCCCAAGCA 60
QY      305 SerAlaGlySerAspTyrCySasnThrGluAsnGluGluLeuPheLeuIleGlnAsp 324
Db      61 AGTCTGGGTCTGATTACTGCAACACAGAACGAAAGAACTTTCTCATACAGAT 120
QY      325 GlnGlnGlyLeuIleThrLeuGlyTTrpIleHisThrHisProThrGlnThrAlaPheLeu 344
Db      121 CAGCAGGGCCTCATCACTGGGCTGGATTCTATCTACCCCAACAGACCGGCTTCTC 180
QY      345 SerSerValAspLeuHisThrHisCySserTyrGlnMetLeuProGlnSerValAla 364
Db      181 TCCAGTGTGCAGCTACACACTGCTCTTACCAAGATGATGTGCCAGAGTCAGTAGCC 240
QY      365 IleValCySerProLysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeu 384
Db      241 ATTGTTGCTCCCAAGTTCAGGAATCGATCTTTAACTAATGACATGACATGAGACTA 300
QY      385 GlnGluIleSerSerCySArgGlnLysGlyPheHisProHisSerLysAspPropoleu 404
Db      301 GAGGAGATTCTCTCTGCTCCCGAAGAAAGATTTCATCCACAGCAAGATCCACTCTG 360
QY      405 PheCySerCySeSerHisValThrValAlaSPaRgaIaValThrIleThraSPLeuArg 424
Db      361 TTCGTGAGTGCAGCCACGTAAGTGTGTGGACAGACGAGTGCATCAAGACCTTCGA 420

RESULT 5
US-09-216-003A-18
; Sequence 18, Application US/09216003A
; Patent No. 6670463
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
; FILE REFERENCE: 210121.462
; CURRENT APPLICATION NUMBER: US/09/216, 003A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1041

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (344)
; OTHER INFORMATION: where n is a, c, g or t
US-09-216-003A-18

```

```

Alignment Scores:
Pred. No.: 3 698-77 1041
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0
DB: 4 Gaps: 0

```

US-09-831-452-1 (1-424) x US-09-216-003A-18 (1-1041)

```

QY 285 LeuCySGlyLysLeuMeTArGAsnGluPheThrIleThrHisValLeuIleProLysGln 304
Db 1 CTCGTGGAAGAACTGATGAGGATGAATTACCATTAACCATGTTCTCTATCCCAAGCAA 60
QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGlnGluLeuPheLeuIleGlnAsp 324
Db 61 AGTGTGGGTCTGATTCTGCAACAGAGAACGAAAGAACTTTCTCTATACAGAT 120
QY 325 GlnGlnGlyLeuIleThrLeuGlyTrrPheIleIsthrHisProThrGlnThrAlaPheLeu 344
Db 121 CAGAGGGCTCATACACACTGGGCTGGATTCATACACCCACAGACCGGCTTCTC 180
QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuProGlnSerValAla 364
Db 181 TCCAGTGTGCACTCAGACACTGCTCTTACAGATGATGTGGCAGAGTACAGTACC 240
QY 365 IleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeu 384
Db 241 ATTGTTGCTCTCCCAAGTTCCAGAAACGATTTCTTTAACTAAGTACGATGAGCTA 300
QY 385 GluGlnIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProProLeu 404
Db 301 GAGAGAGTTTCTTCTCTGCGCCAGAAAGATTTCATCCACAGCAGATCCACCTCTG 360
QY 405 PheCysSerCysSerHisValThrValAlaAspArgAlaValThrIleThrAspLeuArg 424
Db 361 TTCTGTAGCTGACACCGACGTCGATCTGTGTGACAGAGCACTGACATCAGACCTTCA 420

```

RESULT 6

US-09-404-879A-19
Sequence 19, Application US/09404879A
Patent No. 6468546

GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.

APPLICANT: King, Gordon E.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

FILE REFERENCE: 210121.462C2

CURRENT APPLICATION NUMBER: US/09/404,879A

NUMBER OF SEQ ID NOS: 393

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 19

LENGTH: 1043

TYPE: DNA

ORGANISM: Homo sapien

US-09-404-879A-19

Alignment Scores:

```

Pred. No.: 3 71e-77 1043
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0

```

DB: 4 Gaps: 0

US-09-831-452-1 (1-424) x US-09-404-879A-19 (1-1043)

```

QY 285 LeuCySGlyLysLeuMeTArGAsnGluPheThrIleThrHisValLeuIleProLysGln 304
Db 1 CTCGTGGAAGAACTGATGAGGATGAATTACCATTAACCATGTTCTCTATCCCAAGCAA 60
QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGlnGluLeuPheLeuIleGlnAsp 324
Db 61 AGTGTGGGTCTGATTCTGCAACAGAGAACGAAAGAACTTTCTCTATACAGAT 120
QY 325 GlnGlnGlyLeuIleThrLeuGlyTrrPheIleIsthrHisProThrGlnThrAlaPheLeu 344
Db 121 CAGAGGGCTCATACACACTGGGCTGGATTCATACACCCACAGACCGGCTTCTC 180
QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuProGlnSerValAla 364
Db 181 TCCAGTGTGCACTCAGACACTGCTCTTACAGATGATGTGGCAGAGTACAGTACC 240
QY 365 IleValCysSerProLysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeu 384
Db 241 ATTGTTGCTCTCCCAAGTTCCAGAAACGATTTCTTTAACTAAGTACGATGAGCTA 300
QY 385 GluGlnIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProProLeu 404
Db 301 GAGAGAGTTTCTTCTCTGCGCCAGAAAGATTTCATCCACAGCAGATCCACCTCTG 360
QY 405 PheCysSerCysSerHisValThrValAlaAspArgAlaValThrIleThrAspLeuArg 424
Db 361 TTCTGTAGCTGACACCGACGTCGATCTGTGTGACAGAGCACTGACATCAGACCTTCA 420

```

RESULT 7

US-09-338-933-19
Sequence 19, Application US/09338933
Patent No. 6488931

GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer Lynn

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF

FILE REFERENCE: 210121.462C1

CURRENT APPLICATION NUMBER: US/09/338,933

NUMBER OF SEQ ID NOS: 312

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 19

LENGTH: 1043

TYPE: DNA

ORGANISM: Homo sapien

US-09-338-933-19

Alignment Scores:

```

Pred. No.: 3 71e-77 1043
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0
DB: 4 Gaps: 0

```

US-09-831-452-1 (1-424) x US-09-338-933-19 (1-1043)

```

QY 285 LeuCySGlyLysLeuMeTArGAsnGluPheThrIleThrHisValLeuIleProLysGln 304
Db 1 CTCGTGGAAGAACTGATGAGGATGAATTACCATTAACCATGTTCTCTATCCCAAGCAA 60
QY 305 SerAlaGlySerAspTyrCysAsnThrGluAsnGluGlnGluLeuPheLeuIleGlnAsp 324
Db 61 AGTGTGGGTCTGATTCTGCAACAGAGAACGAAAGAACTTTCTCTATACAGAT 120
QY 325 GlnGlnGlyLeuIleThrLeuGlyTrrPheIleIsthrHisProThrGlnThrAlaPheLeu 344
Db 121 CAGAGGGCTCATACACACTGGGCTGGATTCATACACCCACAGACCGGCTTCTC 180

```

QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuProGlnSerValAla 364
|
Db 181 TCCAGTGTGACCTACACACTGCTGCTTACCAAGATGATGTTGCCAGACTCGATGCC 240
QY 365 TLeValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuThrAspHisGlyLeu 384
|
Db 241 ATTGTTGCTCCGCCCAAGTTCCAGAAACTGATCTTTAAACTAATGACATGACATGACTA 300
QY 385 GlnGlnIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProProLeu 404
|
Db 301 GAGGAGATTCTTCTGTCGCGAGAAAGATTTCATCCACACAGCAAGATCCACTCTG 360
QY 405 PheCysSerCysSerHisValThrValValAspArgAlaValThrIleThrAspLeuArg 424
|
Db 361 TTCTGTAGCTGACGACGCTGCTGTGTGACAGACGATGACCATCAGACCTTCCA 420

RESULT 8
US-09-215-681-19
Sequence 19, Application US/09215681A
Patent No. 6528253
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Frudakis, Tony N.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
FILE REFERENCE: 210121.463
CURRENT APPLICATION NUMBER: US/09/215,681A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 310
SOFTWARE: PasteSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1043
TYPE: DNA
ORGANISM: Homo sapien
US-09-215-681-19

Alignment Scores:
Pred. No.: 3,71e-77 Length: 1043
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0
Gaps: 0

US-09-831-452-1 (1-424) x US-09-215-681-19 (1-1043)

QY 285 LeuCysGlyLysLeuMetArgAsnGlnPheThrIleThrHisValLeuIleProLysGln 304
|
Db 1 CTCGTGGAAGAACTGATGAGGAAATGATTTACCATTCACCATGTTCTCATCCCAAGCA 60
QY 305 SerAlaGlySerAspTyrCysAsnThrGlnAsnGlnGlnLeuPheLeuIleGlnAsp 324
|
Db 61 AGTCTGGGTCTGATTACTGCAACACAGAGAAACGAAAGAACTTTCTCATACAGAT 120
QY 325 GlnGlnIleLeuIleThrLeuGlnIleThrIleHisProThrGlnThrAlaPheLeu 344
|
Db 121 CAGCAGGGCTCATCACTGGGTGATTCATCACTCCACACAGACCGGTTTCTC 180
QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuProGlnSerValAla 364
|
Db 181 TCCAGTGTGACCTACACACTGCTGCTTACCAAGATGATGTTGCCAGAGTCAATAGCC 240
QY 365 TLeValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuThrAspHisGlyLeu 384
|
Db 241 ATTGTTGCTCCGCCCAAGTTCCAGAAACTGATCTTTAAACTAATGACATGACATGACTA 300
QY 385 GlnGlnIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProProLeu 404
|
Db 301 GAGGAGATTCTTCTGTCGCGAGAAAGATTTCATCACTCCACAGCAAGATCCACTCTG 360
QY 405 PheCysSerCysSerHisValThrValValAspArgAlaValThrIleThrAspLeuArg 424

Db 361 TTCTGTAGCTGACGACGCTGACTGTGTGTGACAGACGATGACCATCAGACCTTCCA 420

RESULT 9
US-09-216-003A-19

Sequence 19, Application US/09216003A
Patent No. 6670463
GENERAL INFORMATION:

APPLICANT: Mitcham, Jennifer L.
APPLICANT: Frudakis, Tony N.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF OVARIAN CANCER
FILE REFERENCE: 210121.462
CURRENT APPLICATION NUMBER: US/09/216,003A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 310
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 19
LENGTH: 1043
TYPE: DNA
ORGANISM: Homo sapiens
US-09-216-003A-19

Alignment Scores:
Pred. No.: 3,71e-77 Length: 1043
Score: 758.00 Matches: 140
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 34.33% Indels: 0
Gaps: 0

US-09-831-452-1 (1-424) x US-09-216-003A-19 (1-1043)

QY 285 LeuCysGlyLysLeuMetArgAsnGlnPheThrIleThrHisValLeuIleProLysGln 304
|
Db 1 CTCGTGGAAGAACTGATGAGGAAATGATTTACCATTCACCATGTTCTCATCCCAAGCA 60
QY 305 SerAlaGlySerAspTyrCysAsnThrGlnAsnGlnGlnLeuPheLeuIleGlnAsp 324
|
Db 61 AGTCTGGGTCTGATTACTGCAACACAGAGAAACGAAAGAACTTTCTCATACAGAT 120
QY 325 GlnGlnIleLeuIleThrLeuGlnIleThrIleHisProThrGlnThrAlaPheLeu 344
|
Db 121 CAGCAGGGCTCATCACTGGGTGATTCATCACTCCACACAGACCGGTTTCTC 180
QY 345 SerSerValAspLeuHisThrHisCysSerTyrGlnMetMetLeuProGlnSerValAla 364
|
Db 181 TCCAGTGTGACCTACACACTGCTGCTTACCAAGATGATGTTGCCAGAGTCAATAGCC 240
QY 365 TLeValCysSerProLysPheGlnGlnThrGlyPhePheLysLeuThrAspHisGlyLeu 384
|
Db 241 ATTGTTGCTCCGCCCAAGTTCCAGAAACTGATCTTTAAACTAATGACATGACATGACTA 300
QY 385 GlnGlnIleSerSerCysArgGlnLysGlyPheHisProHisSerLysAspProProLeu 404
|
Db 301 GAGGAGATTCTTCTGTCGCGAGAAAGATTTCATCACTCCACAGCAAGATCCACTCTG 360
QY 405 PheCysSerCysSerHisValThrValValAspArgAlaValThrIleThrAspLeuArg 424
|
Db 361 TTCTGTAGCTGACGACGCTGACTGTGTGTGACAGACGATGACCATCAGACCTTCCA 420

RESULT 10
US-09-651-011A-3

Sequence 3, Application US/09651011A
Patent No. 6346416
GENERAL INFORMATION:

APPLICANT: Nicholas M. Dean
APPLICANT: Lex M. Cowart
TITLE OF INVENTION: ANTISENSE MODULATION OF HPK/GCK-LIKE KINASE EXPRESSION
FILE REFERENCE: RTS-0168
CURRENT APPLICATION NUMBER: US/09/651,011A
CURRENT FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 49

```
QY      230   PFCFQVALVALASPATGSEILEUYSRFGQVYAALANSESRASINSEIGLUSEI-----247
          |||||         :|||
Db      1576 CCACCAATGGCTCCCGCATCA--GAGTCTTTTCCTCAATGGCAACTCCGAATCTGTGCAT*1533
          |||||         :|||
QY      248   -----ILPGRTHRLAASPGLYLEARGHLEVAL257
          :|||
```

QY 143 Gluysglnysglnargvalalaglnglnysglnnglnglnleuglnglnluphe 162
::: ::::: ::::: ::::: ||| ::::: |||
DB 1729 GTAGAGGTTGAAGAAATCTACCCGAGATGATAAAATGMAAAAAAGTTCAACATGAATACTA 1788

```

QY 163 HisAlaPheGlu-----GluMetIleArgAsnGlnGluLeu 174
Db 1789 GAGGTGAGAAATCTTCCAGAAATGTTGAATGTGAAGAGACCATCCAAACAAT 1848
QY 175 GluYsgIuArgLeuYsIleValGln-----GluPheGlyLysVal 188
Db 1849 AACATGAAATATGAACTATTAACCCAGAGAAAAAAGAAATGATTAGT----- 1902
QY 189 AspProGlyLeuGlyGlyProLeuValProAspLeuGlyLysProSerLeuAspValPhe 208
Db 1903 -----CTTGAGAGAAAAGCAATTCCACAGAACCCGTGTGA 1938
QY 209 ProThrIleuThrValSer-----SerIleGlnProSerAspCysHisThrThrVal 225
Db 1939 CCTACATTAATGAAATGAAACGTTACTCCCAACCATCTGAAAGGTGAATCCACT--- 1995
QY 226 ArgProAlaLysProProValValAspArgSerIleuLysProGlyAlaLeuSerAsnSer 245
Db 1996 -----AAACCAAGATATATGTTCAAAATTAATAGTACAGAAATTAACCAATTA 2046
QY 246 GluSerIleProThrIleAspGlyLeuArgHisVal 257
Db 2047 AAGGAAACACCAAGATATGATGATGTCCAAAACATGTA 2082

RESULT 12
US-09-566-921-42
; Sequence 42, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debra W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 42
; LENGTH: 6856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 401746.2
US-09-566-921-42

Alignment Scores:
Pred. No.: 0.00358 Length: 6856
Score: 125.00 Matches: 59
Percent Similarity: 35.58% Conservative: 57
Best Local Similarity: 18.10% Mismatches: 108
Query Match: 5.66% Indels: 102
DB: Gaps: 10

US-09-831-452-1 (1-424) x US-09-566-921-42 (1-6856)
QY 42 GluIleIleArgMetAlaSerIleIleYserGluGluGlyAsnIleGluHisAlaPheIle 61
Db 3470 GAAAGGCTGAGATATGAACTCTGAAAGAGTTCAGTCGCTGAGAGAAAGATATGAT 3529
QY 62 LeuYrAsnLysIleThrLeuPheIleGlu-----LysLeuProLysHisArg----- 78
Db 3530 GACCTCAAGGAAAGATGACCTTATGTCATGTCCTCAAGCTCGACACAAAGAGAA 3589
QY 79 -----AspTyr-----LysSerAlaValIleProGlyLys 88
Db 3590 GACTCCACCCACAGACAGCAAGCTGTGATATATCTTTAGCTGAAATTTGAGAAATG 3649
QY 89 LysAspThrValLysLysLeuLysGluIleAlaPheProLysIleGluGluLeuLysAla 108
Db 3650 GAAGACATTCATCAACAGACAGGAAACCAAGT---GAAGAAGAGTACTCTTGACATG 3706

```

```

QY 109 GluLeuLeuLysArgThrThrLysGluTyrThrGluTyrAsnGlnGluLysLysGlu 128
Db 3707 TCATTTCTCTTTAACTCCAGAGCGGTCACAGAGCTGACGAGAGAGAGAGAGAGAG 3760
QY 129 AlaGluGluLeuAlaArgAsnMetAlaIleGlnGlnGluLeuGluLysGluLysArg 148
Db 3761 -----GTGATTCAGAGATGAGCTGAGACCCGCAAGAGAGAGAGAGAGAGAG 3796
QY 149 ValAlaGlnGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 166
Db 3797 GTGCTCCGAGCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3856
QY 167 -----GluMetIleArgAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 185
Db 3857 TATGATCTACTCAAGCGCTCAAGAACTTGAATCAGAAACAAACAAACAAACAAACAAAC 3916
QY 186 GlyLysValAspProGlyLeuGlyGlyProLeuValProAspLeuGlyLysProSerLeu 205
Db 3917 AATGAGTTCCGCAAGGCCCTCAGTGAAGAAAGTGCCTCCAGAGGTGACGCCAGGTGCA 3976
QY 206 AspValPheProThrLeu-----ThrValSer----- 214
Db 3977 CTGCTCTACTGTGTCTCTCATGAGACAGCTGACCTCTGTGAGCGAGAGAGCTGATCCGC 4036
QY 215 -----SerIleGlnPro 218
Db 4037 AAGAGAGAGCTCCTCATCTTAAGTCTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4096
QY 219 SerAspCysHisThrThrValArgProAlaLysProProValValAspArgSerLeuLys 238
Db 4097 AAGGATGACACAGATACATG----- 4117
QY 239 ProGlyAlaLeuSerAsnSerIleProThrIleAspGlyLeuArgHisValVal 258
Db 4118 -----ACAGATTCACAACTCTTTGGAAGATGTACAAATAATGAAA 4159
QY 259 ValProGlyArgLeuCysProGlnPheLeuGlnIleAlaSerAlaAsnThrAlaArgGly 278
Db 4160 GATTAAGGTGAATATGACACAGACATACATGTTGTTGAAGAAACAAAT----- 4207
QY 279 ValGluThrCysGlyIleLeuCysGlyLysLeuMetArgAsnGluPheThrIleThrHis 298
Db 4207 ----- 4207
QY 299 ValIleuIleProLysGlnSerAlaGlySerAspTyrCysAsnThrGluAsnGluGlu 318
Db 4208 -----AGATCATCTGCTCTGATTAACATGATGATGATGATGATGATGATGATGAT 4252
QY 319 LeuPheLeuIleGlnAsp 324
Db 4253 CTGTGCTGTGTTATGAA 4270

RESULT 13
US-09-688-188B-9
; Sequence 9, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3798

```



```

QY 98 ILeAlaPheProLysAlaGluLeuLysAlaGluLeuLeuLysArgTyrThrLysGlu 117
Db 1225 CGACGGCTAGAGAGCAACAAGAGAGAGCGGAGAGCTGAGAGAGAGAGAGAGAGAGAG 1284
QY 118 TyrThrGluTyrAsnGluGluLysLysLysLysLysLysLysLysLysLysLysLys 137
Db 1285 CAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1344
QY 138 IleGlnGlnGlu-----LeuGluLysGluLysGluLysGluLysGluLysGlu 154
Db 1345 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1404
QY 155 -----GlnGlnLeuGluGlnGlnGlnPheHisAlaPheGluGluMetIleArgAsn 171
Db 1405 TATATCAGGCGACGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 172 GlnGluLeuGluLysGluLysGluLysLysLysLysLysLysLysLysLysLysLys 191
Db 1441 -----TTGAAAGCTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
QY 192 LeuGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 209
Db 1485 ACTGAGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1544
QY 210 ThrLeuThrValSerSerIleGlnProSerAspCysHisThrThrValArgProAlaLys 229
Db 1545 GTTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1586
QY 230 ProProValValAspArgSerLeuLysProGlyValLeuSerAsnSerGlnSerIlePro 249
Db 1587 CCA-----TAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1616
QY 250 ThrIleAspGlyLeuArgHisValValValProGlyArgLysCysProGlnPheLeuGln 269
Db 1616 ----- 1616
QY 270 LeuAlaSerAlaAsnThrAlaArgGlyValGluThrCysGlyIleLeuLysGlyLysLeu 289
Db 1617 ---GCAACCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1673
QY 290 MetArgAsnGluPheThrIleThrHisValLeuIleProLysGlnSerAlaGlySerAsp 309
Db 1674 CAAGC-----CCACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1691
QY 310 TyrCysAsnThrGluAsnGluGluGluLeuPheLeuIleGlnAspGlnGlnGlnGln 329
Db 1692 ---TGTG-AACGAGCGGAG-----AGGTGAAGATA 1720
QY 330 ThrLeuGlyTyrIleHisThrHisProThrGlnThrAlaPheLeuSerSerValAspLeu 349
Db 1721 GATTAGAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1780
QY 350 HisThrHis-----CysSerTyrGlnMetLeuProGluSerValAlaIleValCys 367
Db 1781 TGAGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840
QY 368 SerPro-----LysPheGlnGluThrGlyPhePheLysLeuThrAspHisGlyLeuGlu 385
Db 1841 ATCCCGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
QY 386 -----GluIleSerSerCysArgGlnLysGlyPheHisProHisSerLys 400
Db 1901 CTGTTCCTGCTCCGCTGAGATTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1960
QY 401 Asp-----ProProLeu-----PheCys 406
Db 1961 GACGAGAAACTCCACCAAGTATGAGCCAGGCTTCTGT 1999

```

```

/ APPLICANT: Fu, C. Alan
/ TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
/ TITLE OF INVENTION: METHODS OF USE
/ FILE REFERENCE: A-68344/RMS/DHR
/ CURRENT APPLICATION NUMBER: US/09/645,456A
/ CURRENT FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US/09/425,324
/ PRIOR FILING DATE: 1999-10-21
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 6
/ LENGTH: 3972
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: synthetic
/ US-09-645-456A-6

Alignment Scores:
Pred. No.: 0.00925 Length: 3972
Score: 118.00 Matches: 50
Percent Similarity: 47.39% Conservative: 50
Best Local Similarity: 23.70% Mismatches: 73
Query Match: 5.34% Indels: 38
DB: 4 Gaps: 10

US-09-831-452-1 (1-424) x US-09-645-456A-6 (1-3972)
QY 76 LysHisArgAspTyrLysSerAlaValIleProGluLysLysAspThrValLysLysLeu 95
Db 1135 GAGATGAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194
QY 96 LysGlu-----IleAlaPheProLysAlaGluLeuLysAlaGluLeuLysArg 113
Db 1195 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1254
QY 114 TyrThrLysGluTyrThrGluTyrAsnGluLysLysLysLysLysLysLysLysLys 133
Db 1255 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1311
QY 134 ArgAsnMetAlaIleGlnGlnGluLeuLys-----GluLysGlnArgValAla 150
Db 1312 AGCGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1371
QY 151 GlnGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 170
Db 1372 GAAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431
QY 171 AsnGlnGlu-----LeuGluLysGluArgLeu----- 179
Db 1432 CGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1491
QY 180 -----LysIleValGlnGluPheGlyLysValAsp----- 189
Db 1492 CCTAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1551
QY 190 -----ProGlyLysGlyLysProLeuValProAspLeuGluLysProSerLeuAspVal 207
Db 1552 AGTTCCCTGCGATCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1611
QY 208 ---PheProThrLeuThrValSerSerIleGlnProSerAspCysHisThrThrValArg 226
Db 1612 AGGTGAGAGTCTTCAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1671
QY 227 ProAlaLysProPro-----ValValAspArgSerLeuLysProGlyValLeu 242
Db 1672 CCAAGTCAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCCCAAGATCC 1728
QY 243 SerAsnSerGlnSerIle-----ProThr 250
Db 1729 ACCGCTCCAGATCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1761

```

Search completed: August 23, 2004, 21:49:49
Job time : 120 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 14:58:09 ; Search time 30 seconds

(Without alignments)
1672.192 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208

Sequence: 1 MSDHGVSLPPEDRVRLSQ.....FCSCHVTVDRAVITDRL 424

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 654962 seqs, 11835481 residues

Total number of hits satisfying chosen parameters: 654962

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCOT_NEW_COMB.pap:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pap:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pap:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pap:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pap:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pap:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	424	US-10-784-004-756	Sequence 756, App
2	2208	100.0	424	US-10-784-004-1100	Sequence 1100, App
3	2208	100.0	424	US-10-170-205E-30458	Sequence 30458, A
4	2208	100.0	424	US-10-170-205E-31635	Sequence 31635, A
5	2118	95.9	410	US-10-170-205E-30459	Sequence 30459, A
6	1913	86.6	424	US-10-784-004-439	Sequence 439, App
7	1913	86.6	424	US-10-784-004-954	Sequence 954, App
8	1239.5	55.7	463	US-10-370-715B-116	Sequence 116, App
9	1229.5	55.7	463	US-10-170-205E-21521	Sequence 21521, A
10	529	24.0	507	US-10-767-701-45289	Sequence 45289, A
11	521.5	23.6	507	US-10-425-115-311202	Sequence 311202, A
12	513.5	23.3	408	US-10-425-115-198593	Sequence 198593, A
13	411.5	18.6	203	US-10-767-701-40264	Sequence 40264, A
14	398.5	18.0	162	US-10-425-115-311201	Sequence 311201, A
15	387	17.5	78	US-10-793-479-6915	Sequence 6915, App
16	350.5	15.9	274	US-10-425-115-346838	Sequence 346838, App
17	344	15.6	152	US-10-425-115-311198	Sequence 311198, A
18	203	9.2	117	US-10-767-701-411925	Sequence 411925, A
19	131.5	6.0	1212	US-10-170-205E-6618	Sequence 6618, App
20	131.5	6.0	1212	US-10-820-583A-6	Sequence 6, App1
21	131.5	6.0	1212	US-60-582-609-1512	Sequence 1512, App
22	130	5.9	200	US-10-767-701-40123	Sequence 40123, A
23	129	5.8	1296	US-10-170-205E-7174	Sequence 7174, App
24	129	5.8	1296	US-60-582-609-1514	Sequence 1514, App
25	129	5.8	1609	US-60-556-841-4181	Sequence 4181, App
26	128	5.8	2819	US-10-170-205E-28234	Sequence 28234, A

27	127	5.8	261	5	US-09-248-796A-19618	Sequence 19618, A
28	127	5.8	1071	1	PCT-US02-39555A-1735	Sequence 1735, App
29	126.5	5.7	3398	6	US-10-170-205E-18080	Sequence 18080, A
30	126.5	5.7	4919	7	US-60-568-845-241	Sequence 241, App
31	126.5	5.7	5314	7	US-60-568-845-240	Sequence 22664, App
32	126	5.7	446	6	US-10-170-205E-22664	Sequence 22664, A
33	125.5	5.7	77	6	US-10-425-115-336996	Sequence 336996, A
34	125	5.7	1233	6	US-10-725-129A-89	Sequence 89, App1
35	125	5.7	1233	6	US-10-820-583A-4	Sequence 4, App1
36	125	5.7	1233	6	PCT-US04-07412-610	Sequence 610, App
37	124.5	5.6	590	1	US-10-389-559-610	Sequence 610, App
38	124.5	5.6	590	6	US-10-170-205E-22882	Sequence 22882, A
39	124.5	5.6	1235	6	US-10-170-205E-6619	Sequence 6619, App
40	124.5	5.6	1235	7	US-60-582-609-1510	Sequence 1510, App
41	124.5	5.6	923	1	PCT-US04-08006-26	Sequence 26, App1
42	123.5	5.6	1118	6	US-10-170-205E-27197	Sequence 27197, A
43	123.5	5.6	1197	7	US-60-556-841-8238	Sequence 8238, App
44	123.5	5.6	1197	7	US-09-248-796A-20582	Sequence 20582, A
45	123	5.6	428	5		

ALIGNMENTS

RESULT 1
US-10-784-004-756
; Sequence 756, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: Patent version 3.2
; SEQ ID NO: 756
; LENGTH: 424
; TYPE: PRT
; ORGANISM: human
US-10-784-004-756

Query Match 100.0%; Score 2208; DB 6; Length 424;

Best Local Similarity 100.0%; Pred. No. 1.3e+16; Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSDHGVSLPPEDRVRLSQLSGAVEVNEIDIPRRYFRSGVLIIMASYSEGNIEHAF	60
DB	1	MSDHGVSLPPEDRVRLSQLSGAVEVNEIDIPRRYFRSGVLIIMASYSEGNIEHAF	60
QY	61	ILYNYITLFIKLPKRDYKSAVPEKDYTKKLEIAFPABELKAELEKRYTKEYTE	120
DB	61	ILYNYITLFIKLPKRDYKSAVPEKDYTKKLEIAFPABELKAELEKRYTKEYTE	120
QY	121	YNEBKKEAEELAPNMAIQOELEKORVAQKQOOLEOEFHAEEMIRNOELKEXRLK	180
DB	121	YNEBKKEAEELAPNMAIQOELEKORVAQKQOOLEOEFHAEEMIRNOELKEXRLK	180
QY	181	IVQEPKVDPGGLVPLEKSPDVEPTLTVSSIQPSDCHTVTPAKPVPVDSLKRG	240
DB	181	IVQEPKVDPGGLVPLEKSPDVEPTLTVSSIQPSDCHTVTPAKPVPVDSLKRG	240
QY	241	ALNSNSITTLGLSHVVPGRGLCOEFLQASANRARGVTCGILCGKMRNEFTITVL	300
DB	241	ALNSNSITTLGLSHVVPGRGLCOEFLQASANRARGVTCGILCGKMRNEFTITVL	300
QY	301	IKQSGSDYCTNEBEELFLIQDOQGLITLGMITHPTQAFSLSSVDLHTSCSYOMLP	360
DB	301	IKQSGSDYCTNEBEELFLIQDOQGLITLGMITHPTQAFSLSSVDLHTSCSYOMLP	360
QY	361	ESVALYCSKFEPTGFFKLTDLHGLEISSCRKGHPHSKDPPLFCSCHVTVDRAVTI	420
DB	361	ESVALYCSKFEPTGFFKLTDLHGLEISSCRKGHPHSKDPPLFCSCHVTVDRAVTI	420

QY 421 TDLR 424
DB 421 TDLR 424

RESULT 2
US-10-784-004-1100
Sequence 1100, Application US/10784004
GENERAL INFORMATION:

APPLICANT: Biogen Idec
TITLE OF INVENTION: Surrogate Markers of Pain
FILE REFERENCE: 08201.6029-00000
CURRENT APPLICATION NUMBER: US/10/784,004
CURRENT FILING DATE: 2004-02-20
NUMBER OF SEQ ID NOS: 1251
SOFTWARE: PatentIn version 3.2

SEQ ID NO 1100
LENGTH: 424
TYPE: PRT
ORGANISM: human
US-10-784-004-1100

Query Match 100.0%; Score 2208; DB 6; Length 424;
Best Local Similarity 100.0%; Pred. No. 1,3e-146;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQGSAAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLALSQGSAAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
QY 61 ILVNYKITLFLFKLPKRDYSAVPEKKDTVKKLKEIAPFAKAEELKAEELKRYTKYTE 120
DB 61 ILVNYKITLFLFKLPKRDYSAVPEKKDTVKKLKEIAPFAKAEELKAEELKRYTKYTE 120
QY 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
DB 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
QY 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
DB 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
QY 181 IVQEFKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
DB 181 IVQEFKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
QY 241 ALSNSESIPTIDGLRHVVVPERLCPQFLQASANTAGVETCGILCGKLMNEFTITHVL 300
DB 241 ALSNSESIPTIDGLRHVVVPERLCPQFLQASANTAGVETCGILCGKLMNEFTITHVL 300
QY 301 IPKQAGSDYCNTENEBELFLIQOQGLITLGMITHPTQTAPFSSVDLHTHCSYQWMLP 360
DB 301 IPKQAGSDYCNTENEBELFLIQOQGLITLGMITHPTQTAPFSSVDLHTHCSYQWMLP 360
QY 361 ESVAIVSPKQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDDRAVTI 420
DB 361 ESVAIVSPKQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDDRAVTI 420
QY 421 TDLR 424
DB 421 TDLR 424

RESULT 3

US-10-170-205E-30458
Sequence 30458, Application US/10170205E
GENERAL INFORMATION:

APPLICANT: ADAMS, Mark
TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
FILE REFERENCE: CLO01381
CURRENT APPLICATION NUMBER: US/10/170,205E
CURRENT FILING DATE: 2002-06-13
NUMBER OF SEQ ID NOS: 40312
SOFTWARE: PatentIn version 3.2

SEQ ID NO 30458
LENGTH: 424

TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-205E-30458

Query Match 100.0%; Score 2208; DB 6; Length 424;
Best Local Similarity 100.0%; Pred. No. 1,3e-146;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQGSAAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLALSQGSAAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
QY 61 ILVNYKITLFLFKLPKRDYSAVPEKKDTVKKLKEIAPFAKAEELKAEELKRYTKYTE 120
DB 61 ILVNYKITLFLFKLPKRDYSAVPEKKDTVKKLKEIAPFAKAEELKAEELKRYTKYTE 120
QY 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
DB 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
QY 181 IVQEFKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
DB 181 IVQEFKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
QY 241 ALSNSESIPTIDGLRHVVVPERLCPQFLQASANTAGVETCGILCGKLMNEFTITHVL 300
DB 241 ALSNSESIPTIDGLRHVVVPERLCPQFLQASANTAGVETCGILCGKLMNEFTITHVL 300
QY 301 IPKQAGSDYCNTENEBELFLIQOQGLITLGMITHPTQTAPFSSVDLHTHCSYQWMLP 360
DB 301 IPKQAGSDYCNTENEBELFLIQOQGLITLGMITHPTQTAPFSSVDLHTHCSYQWMLP 360
QY 361 ESVAIVSPKQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDDRAVTI 420
DB 361 ESVAIVSPKQETGFFKLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDDRAVTI 420
QY 421 TDLR 424
DB 421 TDLR 424

RESULT 4

US-10-170-205E-31635
Sequence 31635, Application US/10170205E
GENERAL INFORMATION:

APPLICANT: ADAMS, Mark
TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
FILE REFERENCE: CLO01381
CURRENT APPLICATION NUMBER: US/10/170,205E
CURRENT FILING DATE: 2002-06-13
NUMBER OF SEQ ID NOS: 40312
SOFTWARE: PatentIn version 3.2

SEQ ID NO 31635
LENGTH: 424
TYPE: PRT
ORGANISM: Homo sapiens
US-10-170-205E-31635

Query Match 100.0%; Score 2208; DB 6; Length 424;
Best Local Similarity 100.0%; Pred. No. 1,3e-146;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQGSAAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLALSQGSAAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
QY 61 ILVNYKITLFLFKLPKRDYSAVPEKKDTVKKLKEIAPFAKAEELKAEELKRYTKYTE 120
DB 61 ILVNYKITLFLFKLPKRDYSAVPEKKDTVKKLKEIAPFAKAEELKAEELKRYTKYTE 120
QY 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180
DB 121 YNEEKKEAEELARMAIQOLEKEKORVAQKQOQLEQEFHAFEMIRNOLEKERLK 180

```
Db      121 YNEKKKEAEELARNMAIQOELEKEKORVAQOQOQLEQEOFHAFEEIMINQOLEKERLK 180
Qy      181 IVQFEGVNDPGLGPPVLDLEKPSLDVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPG 240
Db      181 IVQFEGVNDPGLGPPVLDLEKPSLDVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPG 240
Qy      241 ALSNSESIPTIDGLRHVVPGRLCPQFLQASANTAGVETGGLCKLMRNEFTITHVL 300
Db      241 ALSNSESIPTIDGLRHVVPGRLCPQFLQASANTAGVETGGLCKLMRNEFTITHVL 300
Qy      301 IPKQASGDYCNTENEBELFLIQOQGLITLGMITHPTQTAFSSVDLHTHCSYQWMLP 360
Db      301 IPKQASGDYCNTENEBELFLIQOQGLITLGMITHPTQTAFSSVDLHTHCSYQWMLP 360
Qy      361 ESVAIVGSPKFOETGFPLTDHGLEEISSCRQGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Db      361 ESVAIVGSPKFOETGFPLTDHGLEEISSCRQGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Qy      421 TDLR 424
Db      421 TDLR 424

RESULT 5
US-10-170-205E-30459
; Sequence 30459, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: C0001381
; CURRENT APPLICATION NUMBER: US/10/170, 205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30459
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-30459
```

Query Match Best Local Similarity 100.0%; Pred. No. 2.4e-140; Length 410;

Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MSDHGDVSLPPEDRVRLASQLGSAVEVNEDIPPRRYFRSGVEIIRMAISYSEGNIEHAF 60
Db      1 MSDHGDVSLPPEDRVRLASQLGSAVEVNEDIPPRRYFRSGVEIIRMAISYSEGNIEHAF 60
Qy      61 ILNNKYITLFIKLPKRDYKSAVIPKKDVTYKLEIAFPKAEELKAEILKRYTKEYTE 120
Db      61 ILNNKYITLFIKLPKRDYKSAVIPKKDVTYKLEIAFPKAEELKAEILKRYTKEYTE 120
Qy      121 YNEKKKEAEELARNMAIQOELEKEKORVAQOQOQLEQEOFHAFEEIMINQOLEKERLK 180
Db      121 YNEKKKEAEELARNMAIQOELEKEKORVAQOQOQLEQEOFHAFEEIMINQOLEKERLK 180
Qy      181 IVQFEGVNDPGLGPPVLDLEKPSLDVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPG 240
Db      181 IVQFEGVNDPGLGPPVLDLEKPSLDVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPG 240
Qy      241 ALSNSESIPTIDGLRHVVPGRLCPQFLQASANTAGVETGGLCKLMRNEFTITHVL 300
Db      241 ALSNSESIPTIDGLRHVVPGRLCPQFLQASANTAGVETGGLCKLMRNEFTITHVL 300
Qy      301 IPKQASGDYCNTENEBELFLIQOQGLITLGMITHPTQTAFSSVDLHTHCSYQWMLP 360
Db      301 IPKQASGDYCNTENEBELFLIQOQGLITLGMITHPTQTAFSSVDLHTHCSYQWMLP 360
Qy      361 ESVAIVGSPKFOETGFPLTDHGLEEISSCRQGFHPSKDPPLFC 406
Db      361 ESVAIVGSPKFOETGFPLTDHGLEEISSCRQGFHPSKDPPLFC 406
```

```
RESULT 6
US-10-784-004-439
; Sequence 439, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idex
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 439
; LENGTH: 424
; TYPE: PRT
; ORGANISM: rat
US-10-784-004-439
```

Query Match Best Local Similarity 86.6%; Score 1913; DB 6; Length 424;

Matches 357; Conservative 33; Mismatches 34; Indels 0; Gaps 0;

```
Qy      1 MSDHGDVSLPPEDRVRLASQLGSAVEVNEDIPPRRYFRSGVEIIRMAISYSEGNIEHAF 60
Db      1 MSDHGDVSLPPEDRVRLASQLGSAVEVNEDIPPRRYFRSGVEIIRMAISYSEGNIEHAF 60
Qy      61 ILNNKYITLFIKLPKRDYKSAVIPKKDVTYKLEIAFPKAEELKAEILKRYTKEYTE 120
Db      61 ILNNKYITLFIKLPKRDYKSAVIPKKDVTYKLEIAFPKAEELKAEILKRYTKEYTE 120
Qy      121 YNEKKKEAEELARNMAIQOELEKEKORVAQOQOQLEQEOFHAFEEIMINQOLEKERLK 180
Db      121 YNEKKKEAEELARNMAIQOELEKEKORVAQOQOQLEQEOFHAFEEIMINQOLEKERLK 180
Qy      181 IVQFEGVNDPGLGPPVLDLEKPSLDVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPG 240
Db      181 IVQFEGVNDPGLGPPVLDLEKPSLDVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPG 240
Qy      241 ALSNSESIPTIDGLRHVVPGRLCPQFLQASANTAGVETGGLCKLMRNEFTITHVL 300
Db      241 ALSNSESIPTIDGLRHVVPGRLCPQFLQASANTAGVETGGLCKLMRNEFTITHVL 300
Qy      301 IPKQASGDYCNTENEBELFLIQOQGLITLGMITHPTQTAFSSVDLHTHCSYQWMLP 360
Db      301 IPKQASGDYCNTENEBELFLIQOQGLITLGMITHPTQTAFSSVDLHTHCSYQWMLP 360
Qy      361 ESVAIVGSPKFOETGFPLTDHGLEEISSCRQGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Db      361 ESVAIVGSPKFOETGFPLTDHGLEEISSCRQGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Qy      421 TDLR 424
Db      421 TDLR 424
```

RESULT 7

```
US-10-784-004-954
; Sequence 954, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idex
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 954
; LENGTH: 424
; TYPE: PRT
; ORGANISM: rat
US-10-784-004-954
```

Query Match Best Local Similarity 86.6%; Score 1913; DB 6; Length 424;

Best Local Similarity 84.2%; Pred. No. 5,5e-126; Matches 357; Conservative 33; Mismatches 34; Indels 0; Gaps 0;

```
QY 1 MSDHGVSLPPEDRVRLSOLGSAVENVEDIPPRRYRSGVEIIRMAISYSEGNIEHAF 60
Db 1 MSDHADVSLPPEDRVRLSOLGSAVENVEDIPPRRYRSGVEIIRMAISYSEGNIEHAF 60
QY 61 ILVNYKTLTFLEKLPKRDYSAVPEKKTQVKKLKEIAPKAEELKAEILKRYTKEYE 120
Db 61 ILVNYKTLTFLEKLPKRDYSAVPEKKTQVKKLKEIAPKAEELKAEILKRYTKEYE 120
QY 121 YNEEKKEAEELANNAIIOELEKEKORVAQOQOOLEOFHAFEMIRNOELEKERLK 180
Db 121 YNEEKKEAEELANNAIIOELEKEKORVAQOQOOLEOFHAFEMIRNOELEKERLK 180
QY 121 YNEEKKEAEELANNAIIOELEKEKORVAQOQOOLEOFHAFEMIRNOELEKERLK 180
Db 121 YNEEKKEAEELANNAIIOELEKEKORVAQOQOOLEOFHAFEMIRNOELEKERLK 180
QY 181 IVQFEGVADPGGLGSLVPLDEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRLKXG 240
Db 181 IVQFEGVADPGGLGSLVPLDEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRLKXG 240
QY 241 ALSNSESIPITDGRHVVPVPGRLCPQFLOLASANTARGVETCGILCGKLRNEFTITHYL 300
Db 241 ALSNSENPTIEGRHIVVPRNLCEFLQLASANTARGVETCGVLCGLKLRNEFTITHYL 300
QY 301 IPKOSASDYCNTENEBELFLIOQOGLITLGMITHPTQTALSSVDLTHCSYQMP 360
Db 301 IPKOSASDYCNTENEBELFLIOQOGLITLGMITHPTQTALSSVDLTHCSYQMP 360
QY 361 ESVAIVCSPKFOETGFKLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDAVVI 420
Db 361 ESVAIVCSPKFOETGFKLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDAVVI 420
QY 421 TDLR 424
Db 421 TDLR 424
```

RESULT 8
US-10-370-715B-116
Sequence 116, Application US/10370715B

```
GENERAL INFORMATION:
Applicant: Bodary, Sarah C.
Applicant: Clark, Hillary
Applicant: Brisdell, Hunter
Applicant: Jackson, Janet
Applicant: Schoenfeld, Jill R.
Applicant: Williams, P. Mickey
Applicant: Wood, William I.
Applicant: Wu, Thomas D.
Title of Invention: Compositions and Methods for the Treatment of Immune
Title of Invention: Related Diseases
File Reference: P19481-US
Current Application Number: US/10/370,715B
Current Filing Date: 2003-02-21
Number of SEQ ID NOS: 742
SEQ ID NO 116
Length: 463
Type: PRT
Organism: Homo sapiens
US-10-370-715B-116
```

Query Match 55.7%; Score 1229.5; DB 6; Length 463;
Best Local Similarity 56.4%; Pred. No. 4.2e-78;
Matches 232; Conservative 72; Mismatches 92; Indels 15; Gaps 2;

```
QY 121 YNEEKKEAEELANNAIIOELEKEKORVAQOQOOLEOFHAFEMIRNOELEKERLK 180
Db 121 YNEEKKEAEELANNAIIOELEKEKORVAQOQOOLEOFHAFEMIRNOELEKERLK 180
QY 138 YLQSNKKYKAEILKLEHQLIEARKRIAQMRQOQLESEQFLFEDQLKKQELARQMR 197
Db 138 YLQSNKKYKAEILKLEHQLIEARKRIAQMRQOQLESEQFLFEDQLKKQELARQMR 197
QY 181 IVQFEGVADPGGLGSLVPLDEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVD 234
Db 181 IVQFEGVADPGGLGSLVPLDEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVD 234
QY 198 SQQTSG-----LSQIDGALSICFSSTHONNSLNVFADQPNKSDATVYASHSPVN 248
Db 198 SQQTSG-----LSQIDGALSICFSSTHONNSLNVFADQPNKSDATVYASHSPVN 248
QY 235 RSLKPGALSNSBSIPTIDGLRHVVVPGRLCPQFLOLASANTARGVETCGILCGKLRNEF 294
Db 235 RSLKPGALSNSBSIPTIDGLRHVVVPGRLCPQFLOLASANTARGVETCGILCGKLRNEF 294
QY 249 RALTPAALISAVQNLVWGRLCVLPEDLCHKFQLQLESNVVRGIELTCGILCGKLTNNEF 308
Db 249 RALTPAALISAVQNLVWGRLCVLPEDLCHKFQLQLESNVVRGIELTCGILCGKLTNNEF 308
QY 295 TTHVILPKOSAGSDYCNTENEBELFLIOQOGLITLGMITHPTQTALSSVDLTHCS 354
Db 295 TTHVILPKOSAGSDYCNTENEBELFLIOQOGLITLGMITHPTQTALSSVDLTHCS 354
QY 309 TTHVIVKOSAGSDYCNTENEBELFLIOQOGLITLGMITHPTQTALSSVDLTHCS 368
Db 309 TTHVIVKOSAGSDYCNTENEBELFLIOQOGLITLGMITHPTQTALSSVDLTHCS 368
QY 355 YQMLPESVAIVCSPKFOETGFKLTDHGLEIISCRQKGFHPSKDPPLF 405
Db 355 YQMLPESVAIVCSPKFOETGFKLTDHGLEIISCRQKGFHPSKDPPLF 405
QY 369 YQMLPESVAIVCSPKFOETGFKLTDHGLEIISCRQKGFHPSKDPPLF 419
Db 369 YQMLPESVAIVCSPKFOETGFKLTDHGLEIISCRQKGFHPSKDPPLF 419
```

RESULT 9
US-10-170-205E-21521
Sequence 21521, Application US/10170205E

```
GENERAL INFORMATION:
Applicant: Adams, Mark
Title of Invention: Devices, Such as Arrays, Comprised of Human Proteins or Protein
Title of Invention: Capture Agents, and Uses Thereof
File Reference: CL001381
Current Application Number: US/10/170,205E
Current Filing Date: 2002-06-13
Number of SEQ ID NOS: 40312
Software: PatentIn version 3.2
SEQ ID NO 21521
Length: 463
Type: PRT
Organism: Homo sapiens
US-10-170-205E-21521
```

Query Match 55.7%; Score 1229.5; DB 6; Length 463;
Best Local Similarity 56.4%; Pred. No. 4.2e-78;
Matches 232; Conservative 72; Mismatches 92; Indels 15; Gaps 2;

```

US-10-767-701-45299
; Sequence 45299, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; CURRENT FILING DATE: 2004-01-25
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 45299
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; NAME/KEY: unsure
; LOCATION: (1)..(507)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C9218_1.pap
US-10-767-701-45299

Query Match      24.0%; Score 529; DB 6; Length 507;
Best Local Similarity 31.5%; Pred. No. 4.9e-29;
Matches 163; Conservative 74; Mismatches 164; Indels 116; Gaps 22;

QY 10 PPED-RVALSQTGSA--VEVNEIDIPRRYRSGVEIIRMASIYSEBGNIEHAFILNKY 66
DB 3 PPGARVAGALINIEKCAPRIADVHRIRLPYFRIGSLIRQANIYRNSNLDLVILRLY 62
QY 67 ITFIEKLPKRRDYKSAVIRPKOTVKKLEIA-FPAEELKAEI-LKRYT-KEYTEYN 122
DB 63 SSLCETIPKRDYHAF-----KLKKAFFDKLNDVMEKLETLKRVQROVAEHN 112
QY 123 EEKKEA--EELANMAIQELEKEXKORVNAQKQO-----LEBOQHAFEBEMIRNQ 172
DB 113 RGVVESNTNLSNGNYGTRRIE--ORTPTSYTPQPFVSGTNGALQKTHAGQVALLP 169
QY 173 ELEKRLKI-----VOEFGKVP-GLGG-----PLVP- 198
DB 170 SVQKQFNNLPYKKEETLARSILGPNGLNGRMNGPVTEIKVQYPSNEELSTRDITSVPS 229
QY 199 -----DLEKSLDVFPFLTVSIOPSDCHTV----- 225
DB 230 ILNQDDSHGPGSTAPAPPPDSSNDNDNMRKSVLSLDDGRWSVPAERTPLPSASLEBELFQ 289
QY 226 ----RPAKPPY--VDRSLKPGALSN-SESIPITDGLR--HVVVPGRLCPQFIQLASANT 275
DB 290 LDIDQSPRPVLAEVQPIPSRVADPTPLGTPISGTRAFONLHVPIKMECFLRVASNT 349
QY 276 ARGVETGILICGKIMNEFTITVHLIPKQAGSDYCNTENEBELFLIQOQGLITLGMW 335
DB 350 KRSLETCGVAGTLKATFCVTLIIPKQSTSTNCAINEBELFEVQDGSIFLTGMW 409
QY 336 THPTQTAFLSSVDLHTCSYOMMLPESVAIVCSP--KFGCTGPFKLTLD-HGLEIISCRQ 392
DB 410 THPTQSCXSSIDLHNHYIQVMLEPAIYVAPDTTTRKHGIFHLTDPGGMGVHDCOE 469
QY 393 KGFHPSKDP---PLFCSCSHYTV-VDRAVTITDLR 424
DB 470 RGFHPH-KAPLDGSPYEQCSHYVMNDIKFDMIDLR 505

```

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 311202
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46872C.1.pap
US-10-425-115-311202

Query Match      23.6%; Score 521.5; DB 6; Length 507;
Best Local Similarity 30.3%; Pred. No. 1.6e-28;
Matches 151; Conservative 73; Mismatches 162; Indels 113; Gaps 19;

QY 25 VEVNEIDIPRRYRSGVEIIRMASIYSEBGNIEHAFILNKYITLPIEKLPHGRDYKSAV 84
DB 21 IAVDHRISLPYFRIGSLIRQANIYRNSNLDLVILRLYSSLCEITIPRHRODHAK 80
QY 85 IPEKK-----DTYKLEIAFPKAEELKAEILKRYKTYNEKKKE--AEELANM 136
DB 81 LREKAFEDKLVDIKEL-----ETLKPVV--QROIVHNHNGTVEFTNSINGNY 128
QY 137 AIQOELEK-----EKORVAQKQOQLBOEQHAFEBEMIRNDELKERLKI-----VQ 183
DB 129 GITHREQCTPSSSYTPQTITGVSSNGAL-OKPHAGQVASSSVQKQFNNLPYKAEETIA 187
QY 184 EFGKVP-GLGG-----PLVPD 199
DB 188 RSHILGPNGLNGRMNGPVTVGIRKQYPSNEELTQNDITSLVPSILNQDDLHGSNTAPAPPP 247
QY 200 LEKPSLDVFPFLTVSIOQ-----PSDCHTV-----RAKPPY--VDR 235
DB 248 PSSSSNDNDNMRKSVLSLDDGRWSVPAERTPLPSANLEBELFOLNIKQSPRPVLAEVQR 307
QY 236 SLKPGALSN-SESIPITDGLR--HVVVPGRLCPQFIQLASANTAGVETGILICGLMKN 292
DB 308 PISPSRVADPTPLGTPISGTRAFONLHVPIKMECFLRVASNTKSLDFCGVLAGTLKNR 367
QY 293 EFTITVHLIPKQAGSDYCNTENEBELFLIQOQGLITLGMWHTHPTQTAFLSSVDLHT 352
DB 368 TTYVITLIPKQSTSVYCEATNEBELFEVQDMSLFTIGWHTHPTQSCFLSSVDLHNH 427
QY 353 GSYOMMLPESVAIVCSP--KFGCTGPFKLTLD-HGLEIISCRQKGFHPSKDP---PLFC 406
DB 428 YQYQVMLEPAIYVAPDTTTRKHGIFHLTDPGGMGVHDCOEGRHPH-KAPLDGSPYIK 486
QY 407 SCSHYTV-VDRAVTITDLR 424
DB 487 QCSHYVMDTIKFDIMIDLR 505

```

```

RESULT 11
US-10-425-115-311202
; Sequence 311202, Application US/10425115
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```


5 PPSRVADPRPG-IATSET----GRYONLHVPAIMECFLSVAELIANNLEICGLIAGIL 59

NUMBER OF SEXED IND. 2004

SOFTWARE: Patent.pm
SEQ ID NO 6915
LENGTH: 78
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 31
OTHER INFORMATION: Xaa=Ile or Ser
FEATURE:
NAME/KEY: UNSURE
LOCATION: 33
OTHER INFORMATION: Xaa=Pro or Arg
US-10-793-479-6915

Query Match 17.5%; Score 387; DB 6; Length 78;
Best Local Similarity 97.4%; Pred. No. 4e-20; Mismatches 0; Gaps 0;
Matches 76; Conservative 0; Indels 2;

Oy	1	MSDHGDVSLPPEDRVRAISQLGSAVEVNEDIPPRRYFPGSGVEIIRMAISYSEGNIEHAF	60
Db	1	MSDHGDVSLPPEDRVRAISQLGSAVEVNEDIKXRRYFPGSGVEIIRMAISYSEGNIEHAF	60
Oy	61	ILYNYITLFTFKLPKGR	78
Db	61	ILYNYITLFTFKLPKGR	78

Search completed: August 17, 2004, 15:03:44
Job time : 31 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 14:57:49 ; Search time 177 Seconds
(without alignments)
2338.117 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208
Sequence: 1 MSDHGDVSLPEEDRVRLSQ.....FCSCSHVTVDRAVTTDLR 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents AA Main:*

1: /cgn2_6/prodata/2/paa/US06_COMB.pep:*

2: /cgn2_6/prodata/2/paa/US06_COMB.pep:*

3: /cgn2_6/prodata/2/paa/US07_COMB.pep:*

4: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

5: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

6: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

7: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

8: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

9: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

10: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

11: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

12: /cgn2_6/prodata/2/paa/US08_COMB.pep:*

13: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

14: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

15: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

16: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

17: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

18: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

19: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

20: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

21: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

22: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

23: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

24: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

25: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

26: /cgn2_6/prodata/2/paa/US09_COMB.pep:*

27: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

28: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

29: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

30: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

31: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

32: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

33: /cgn2_6/prodata/2/paa/US10_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	ID	Description
------------	-------	--------------	----	-------------

1	2208	100.0	424	1	PCT-US02-00976-7	Sequence 7, Appli
2	2208	100.0	424	23	US-09-831-452-1	Sequence 1, Appli
3	2208	100.0	424	26	US-10-046-961-7	Sequence 7, Appli
4	2208	100.0	424	26	US-10-047-253-7	Sequence 6, Appli
5	2208	100.0	424	26	US-10-340-578-66	Sequence 6, Appli
6	2208	100.0	424	31	US-10-671-572A-1	Sequence 2, Appli
7	2208	100.0	424	33	US-10-752-442-2	Sequence 1, Appli
8	2208	100.0	424	33	US-60-452-680-24059	Sequence 24059, A
9	2208	100.0	424	33	US-60-452-680-24060	Sequence 24060, A
10	2208	100.0	424	33	US-60-453-050-15016	Sequence 15016, A
11	2208	100.0	424	33	US-60-453-050-15017	Sequence 15017, A
12	2208	100.0	424	33	US-60-453-135-15016	Sequence 15016, A
13	2208	100.0	424	33	US-60-453-135-15017	Sequence 15017, A
14	2208	100.0	424	33	US-60-455-444-8192	Sequence 8192, Ap
15	2208	100.0	424	33	US-60-455-444-8193	Sequence 8193, Ap
16	2208	100.0	424	33	US-60-455-444-8194	Sequence 8194, Ap
17	2208	100.0	424	33	US-60-465-241-8192	Sequence 8192, Ap
18	2208	100.0	424	33	US-60-465-241-8193	Sequence 8193, Ap
19	2208	100.0	424	33	US-60-465-412-15016	Sequence 15016, A
20	2118	95.9	410	33	US-60-465-412-15017	Sequence 15017, A
21	2118	95.9	410	33	US-60-452-680-24061	Sequence 24061, A
22	2118	95.9	410	33	US-60-453-050-15018	Sequence 15018, A
23	2118	95.9	410	33	US-60-453-135-15018	Sequence 15018, A
24	2118	95.9	410	33	US-60-455-444-8194	Sequence 8194, Ap
25	2118	95.9	410	33	US-60-465-241-8194	Sequence 8194, Ap
26	1904	86.2	424	23	US-09-831-452-3	Sequence 15018, A
27	1904	86.2	424	31	US-10-671-572A-3	Sequence 3, Appli
28	1269.5	57.5	421	1	PCT-US02-00976-5	Sequence 5, Appli
29	1269.5	57.5	421	26	US-10-046-961-5	Sequence 5, Appli
30	1269.5	57.5	421	26	US-10-047-253-5	Sequence 5, Appli
31	1269.5	57.5	421	29	US-10-340-578-67	Sequence 6, Appli
32	1229.5	55.7	461	1	PCT-US02-00976-6	Sequence 6, Appli
33	1229.5	55.7	461	18	US-09-488-725A-3556	Sequence 3556, Ap
34	1229.5	55.7	461	26	US-10-046-961-6	Sequence 6, Appli
35	1229.5	55.7	461	26	US-10-047-253-6	Sequence 6, Appli
36	1229.5	55.7	461	28	US-10-258-898A-3556	Sequence 3556, Ap
37	1229.5	55.7	461	28	US-10-266-897-3556	Sequence 3556, Ap
38	1229.5	55.7	461	29	US-10-340-578-68	Sequence 68, Appli
39	1229.5	55.7	463	26	US-10-012-600B-167	Sequence 167, App
40	1229.5	55.7	463	26	US-60-452-680-16473	Sequence 16473, A
41	871.5	39.5	420	20	US-09-614-150-3090	Sequence 3090, Ap
42	871.5	39.5	420	33	US-09-614-150A-3090	Sequence 3090, Ap
43	871.5	39.5	420	33	US-60-191-637-3096	Sequence 3096, Ap
44	871.5	39.5	420	33	US-60-191-637-3096	Sequence 3096, Ap
45	846	38.3	236	22	US-09-760-476-2534	Sequence 2534, Ap

ALIGNMENTS

RESULT 1

PCT-US02-00976-7

Sequence 7, Application PC/TUS0200976

GENERAL INFORMATION:

APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY

APPLICANT: COPE, Gregory

APPLICANT: VERVA, Rati

APPLICANT: ARAVIND, L.

APPLICANT: KOONIN, Eugene V.

TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS

FILE REFERENCE: CITI510-4W0

CURRENT APPLICATION NUMBER: PCT/US02/00976

CURRENT FILING DATE: 2002-01-14

PRIOR APPLICATION NUMBER: US 60/261,314

PRIOR FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: US 60/322,322

PRIOR FILING DATE: 2001-09-14

PRIOR APPLICATION NUMBER: US 60/322,030

PRIOR FILING DATE: 2001-09-14

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

```

; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-us02-00976-7

Query Match      100.0%; Score 2208; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 3,1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60

QY 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120
DB 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120

QY 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180
DB 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180

QY 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180
DB 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180

QY 181 IVQFGKVDPLGGPLVLDLEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRSLKPG 240
DB 181 IVQFGKVDPLGGPLVLDLEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRSLKPG 240

QY 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETCGILGKLMRNEFTITHVL 300
DB 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETCGILGKLMRNEFTITHVL 300

QY 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETCGILGKLMRNEFTITHVL 300
DB 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETCGILGKLMRNEFTITHVL 300

QY 301 IPKOSAGSDYCNTENEBELFLIQOQGLITLGMHTPTQTAFISSVDLTHCSYQMMLP 360
DB 301 IPKOSAGSDYCNTENEBELFLIQOQGLITLGMHTPTQTAFISSVDLTHCSYQMMLP 360

QY 361 ESVAIVCSPKQETGFELTDHGLEISGCRQGFHSHKDPPLFCSCHVTVDRAVTI 420
DB 361 ESVAIVCSPKQETGFELTDHGLEISGCRQGFHSHKDPPLFCSCHVTVDRAVTI 420

QY 421 TDLR 424
DB 421 TDLR 424

RESULT 2
US-09-831-452-1
; Sequence 1, Application US/09831452
; GENERAL INFORMATION:
; APPLICANT: SUGAMURA, Kazuo
; TITLE OF INVENTION: Protein AMSh and CDNA thereof
; FILE REFERENCE: 2001-0572A/LC/00653
; CURRENT APPLICATION NUMBER: US/09/831,452
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/J999/06309
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: JP No. 10-322674
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-831-452-1

Query Match      100.0%; Score 2208; DB 23; Length 424;
Best Local Similarity 100.0%; Pred. No. 3,1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60

QY 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120
DB 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120
```

```

DB 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120
QY 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180
DB 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180

QY 181 IVQFGKVDPLGGPLVLDLEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRSLKPG 240
DB 181 IVQFGKVDPLGGPLVLDLEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRSLKPG 240

QY 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETCGILGKLMRNEFTITHVL 300
DB 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETCGILGKLMRNEFTITHVL 300

QY 301 IPKOSAGSDYCNTENEBELFLIQOQGLITLGMHTPTQTAFISSVDLTHCSYQMMLP 360
DB 301 IPKOSAGSDYCNTENEBELFLIQOQGLITLGMHTPTQTAFISSVDLTHCSYQMMLP 360

QY 361 ESVAIVCSPKQETGFELTDHGLEISGCRQGFHSHKDPPLFCSCHVTVDRAVTI 420
DB 361 ESVAIVCSPKQETGFELTDHGLEISGCRQGFHSHKDPPLFCSCHVTVDRAVTI 420

QY 421 TDLR 424
DB 421 TDLR 424

RESULT 3
US-10-046-961-7
; Sequence 7, Application US/10046961
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: LYAPINA, Svetlana
; APPLICANT: VERMA, Rati
; APPLICANT: DESHAIR, Raymond
; TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS
; FILE REFERENCE: CIT1310-3
; CURRENT APPLICATION NUMBER: US/10/046,961
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-046-961-7

Query Match      100.0%; Score 2208; DB 26; Length 424;
Best Local Similarity 100.0%; Pred. No. 3,1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMAISYSEGNIEHAF 60

QY 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120
DB 61 ILVNYKITLFIETKLPKRDYKSAVIPKQDVTKKLKEIAFPKAEELKAEILKRYTKEYTE 120

QY 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180
DB 121 YNEEKKEAEELARNAVAIQOLEKEKORVAQOQOQLEQOFHAFEEIMRNQELERK 180

QY 181 IVQFGKVDPLGGPLVLDLEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRSLKPG 240
DB 181 IVQFGKVDPLGGPLVLDLEKPSLDVFPFLTVSSIQPSDCHTTVPAPKPVVDRSLKPG 240
```

QY 241 ALSNSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKLMRNEFTITHVL 300
Db 241 ALSNSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKLMRNEFTITHVL 300
QY 301 IPKQASGDYCNTENEEELFLIQDQGLITLGMHTHTPTOTAFILSSVDLHTHCSYQMMLP 360
Db 301 IPKQASGDYCNTENEEELFLIQDQGLITLGMHTHTPTOTAFILSSVDLHTHCSYQMMLP 360
QY 361 ESVAIVCSPKFOETGFFLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
Db 361 ESVAIVCSPKFOETGFFLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
QY 421 TDLR 424
Db 421 TDLR 424

RESULT 4
US-10-047-253-7
; Sequence 7, Application US/10047253
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: COPE, Gregory
; APPLICANT: VERMA, Rati
; APPLICANT: ARAVIND, L
; APPLICANT: KOONIN, Eugene
; APPLICANT: DESHAIES, Raymond
; TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS
; FILE REFERENCE: CIT1510-4
; CURRENT APPLICATION NUMBER: US/10/047,253
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-047-253-7

Query Match 100.0%; Score 2208; DB 26; Length 424;
Best Local Similarity 100.0%; Pred. No. 3,1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDHGDVSLPPEDRVRLSQLGSAVEVNEDIIPRRYFSGVEIIRMAIYSEEGNIEHAF 60
Db 1 MSDHGDVSLPPEDRVRLSQLGSAVEVNEDIIPRRYFSGVEIIRMAIYSEEGNIEHAF 60
QY 61 ILNNKYITLLEKLPKRHDYKSAVIPKCDTVKKLKEIAFPKAEELKAEILKRYTKEYTE 120
Db 61 ILNNKYITLLEKLPKRHDYKSAVIPKCDTVKKLKEIAFPKAEELKAEILKRYTKEYTE 120
QY 121 YNEEKKEAEELARMAIIOQLLEKEKORVAQOQKQOLEQOEFHAFEMIRNOLEKERLX 180
Db 121 YNEEKKEAEELARMAIIOQLLEKEKORVAQOQKQOLEQOEFHAFEMIRNOLEKERLX 180
QY 181 IVOEFGKVDPLGSLVLDLEKPSLDVFPPLTVSSIQSDCHTTVPAPKPVVDRSLKPG 240
Db 181 IVOEFGKVDPLGSLVLDLEKPSLDVFPPLTVSSIQSDCHTTVPAPKPVVDRSLKPG 240
QY 241 ALSNSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKLMRNEFTITHVL 300
Db 241 ALSNSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKLMRNEFTITHVL 300
QY 301 IPKQASGDYCNTENEEELFLIQDQGLITLGMHTHTPTOTAFILSSVDLHTHCSYQMMLP 360
Db 301 IPKQASGDYCNTENEEELFLIQDQGLITLGMHTHTPTOTAFILSSVDLHTHCSYQMMLP 360

QY 361 ESVAIVCSPKFOETGFFLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
Db 361 ESVAIVCSPKFOETGFFLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
QY 421 TDLR 424
Db 421 TDLR 424

RESULT 5
US-10-340-578-66
; Sequence 66, Application US/10340578
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: DESHAIES, Raymond J.
; APPLICANT: COPE, Gregory
; APPLICANT: VERMA, Rati
; APPLICANT: AMEROGGIO, Xavier I.
; TITLE OF INVENTION: MODULATION OF COP9 SIGNALSOME ISOPEPTIDASE ACTIVITY
; FILE REFERENCE: CIT1590-1
; CURRENT APPLICATION NUMBER: US/10/340,578
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/355,334
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 10/047,253
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 10/046,961
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-340-578-66

Query Match 100.0%; Score 2208; DB 29; Length 424;
Best Local Similarity 100.0%; Pred. No. 3,1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDHGDVSLPPEDRVRLSQLGSAVEVNEDIIPRRYFSGVEIIRMAIYSEEGNIEHAF 60
Db 1 MSDHGDVSLPPEDRVRLSQLGSAVEVNEDIIPRRYFSGVEIIRMAIYSEEGNIEHAF 60
QY 61 ILNNKYITLLEKLPKRHDYKSAVIPKCDTVKKLKEIAFPKAEELKAEILKRYTKEYTE 120
Db 61 ILNNKYITLLEKLPKRHDYKSAVIPKCDTVKKLKEIAFPKAEELKAEILKRYTKEYTE 120
QY 121 YNEEKKEAEELARMAIIOQLLEKEKORVAQOQKQOLEQOEFHAFEMIRNOLEKERLX 180
Db 121 YNEEKKEAEELARMAIIOQLLEKEKORVAQOQKQOLEQOEFHAFEMIRNOLEKERLX 180
QY 181 IVOEFGKVDPLGSLVLDLEKPSLDVFPPLTVSSIQSDCHTTVPAPKPVVDRSLKPG 240
Db 181 IVOEFGKVDPLGSLVLDLEKPSLDVFPPLTVSSIQSDCHTTVPAPKPVVDRSLKPG 240
QY 241 ALSNSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKLMRNEFTITHVL 300
Db 241 ALSNSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKLMRNEFTITHVL 300
QY 301 IPKQASGDYCNTENEEELFLIQDQGLITLGMHTHTPTOTAFILSSVDLHTHCSYQMMLP 360
Db 301 IPKQASGDYCNTENEEELFLIQDQGLITLGMHTHTPTOTAFILSSVDLHTHCSYQMMLP 360
QY 361 ESVAIVCSPKFOETGFFLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
Db 361 ESVAIVCSPKFOETGFFLTDHGLEEISSCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420

Db 361 ESVAIVCSPKFOETGFEKLTIDHGLEIBISSCRQGFHPSKDPFLFCSCHVTVDRAVTI 420
Qy 421 TDLR 424
Db 421 TDLR 424

RESULT 6

US-10-671-572A-1
Sequence 1, Application US/10671572A
GENERAL INFORMATION:
APPLICANT: SUGAMURA, Kazuo
TITLE OF INVENTION: Protein AMSH and cDNA thereof
FILE REFERENCE: 2003-1382/WMC/00653
CURRENT APPLICATION NUMBER: US/10/671,572A
CURRENT FILING DATE: 2003-09-29
PRIOR APPLICATION NUMBER: 09/831,452
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/JP99/06309
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: JP No. 10-322674
PRIOR FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 424
TYPE: PRT
ORGANISM: Homo sapiens
US-10-671-572A-1

Query Match 100.0%; Score 2208; DB 31; Length 424;
Best Local Similarity 100.0%; Pred. No. 3, 1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDHGDVSLPPEDRVRAISQUGSAVEVNEDIIPRRYFRSGVEIIRMASIYSEEGNIEHAF 60
Db 1 MSDHGDVSLPPEDRVRAISQUGSAVEVNEDIIPRRYFRSGVEIIRMASIYSEEGNIEHAF 60
Qy 61 ILVNYKTYTLFTEKLPKRDYSAVPEKQDVKKLKEIAPFKAEELKAEELKRYTKYTE 120
Db 61 ILVNYKTYTLFTEKLPKRDYSAVPEKQDVKKLKEIAPFKAEELKAEELKRYTKYTE 120
Qy 121 YNEEKKEAEELANMAIQOLEKEKORVAQOQOQOEQFHAPEEMIRNOELEKERLK 180
Db 121 YNEEKKEAEELANMAIQOLEKEKORVAQOQOQOEQFHAPEEMIRNOELEKERLK 180
Qy 181 IVQEPGKVDPLGSPVLDLEKESLDVFPFLLTVSSIQPSDCHTTVPRAKPPVVDRLKPG 240
Db 181 IVQEPGKVDPLGSPVLDLEKESLDVFPFLLTVSSIQPSDCHTTVPRAKPPVVDRLKPG 240
Qy 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTAGVETCGILCKLMRNEFTITHVL 300
Db 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTAGVETCGILCKLMRNEFTITHVL 300
Qy 301 IPKQAGSDYCNTENEEELFLIDQOGLITLGMITHPTQAFLLSSVDLHTHCYQMMLP 360
Db 301 IPKQAGSDYCNTENEEELFLIDQOGLITLGMITHPTQAFLLSSVDLHTHCYQMMLP 360
Qy 361 ESVAIVCSPKFOETGFEKLTIDHGLEIBISSCRQGFHPSKDPFLFCSCHVTVDRAVTI 420
Db 361 ESVAIVCSPKFOETGFEKLTIDHGLEIBISSCRQGFHPSKDPFLFCSCHVTVDRAVTI 420
Qy 421 TDLR 424
Db 421 TDLR 424

RESULT 7
US-10-752-442-2
Sequence 2, Application US/10752442
GENERAL INFORMATION:
APPLICANT: Itoh, Fumiko

APPLICANT: Itoh, Susumu
APPLICANT: Heidin, Carl-Henrik
APPLICANT: ten-Dijke, Peter
TITLE OF INVENTION: SMAD ASSOCIATING POLYPEPTIDES
FILE REFERENCE: L0461.70096US01
CURRENT APPLICATION NUMBER: US/10/752,442
CURRENT FILING DATE: 2004-01-06
PRIOR APPLICATION NUMBER: US 09/665,479
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US 60/154,846
PRIOR FILING DATE: 1999-09-20
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 424
TYPE: PRT
ORGANISM: Homo sapiens
US-10-752-442-2

Query Match 100.0%; Score 2208; DB 32; Length 424;
Best Local Similarity 100.0%; Pred. No. 3, 1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSDHGDVSLPPEDRVRAISQUGSAVEVNEDIIPRRYFRSGVEIIRMASIYSEEGNIEHAF 60
Db 1 MSDHGDVSLPPEDRVRAISQUGSAVEVNEDIIPRRYFRSGVEIIRMASIYSEEGNIEHAF 60
Qy 61 ILVNYKTYTLFTEKLPKRDYSAVPEKQDVKKLKEIAPFKAEELKAEELKRYTKYTE 120
Db 61 ILVNYKTYTLFTEKLPKRDYSAVPEKQDVKKLKEIAPFKAEELKAEELKRYTKYTE 120
Qy 121 YNEEKKEAEELANMAIQOLEKEKORVAQOQOQOEQFHAPEEMIRNOELEKERLK 180
Db 121 YNEEKKEAEELANMAIQOLEKEKORVAQOQOQOEQFHAPEEMIRNOELEKERLK 180
Qy 181 IVQEPGKVDPLGSPVLDLEKESLDVFPFLLTVSSIQPSDCHTTVPRAKPPVVDRLKPG 240
Db 181 IVQEPGKVDPLGSPVLDLEKESLDVFPFLLTVSSIQPSDCHTTVPRAKPPVVDRLKPG 240
Qy 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTAGVETCGILCKLMRNEFTITHVL 300
Db 241 ALSNSESIPITDGLRHVVVPGRLCPQFLQASANTAGVETCGILCKLMRNEFTITHVL 300
Qy 301 IPKQAGSDYCNTENEEELFLIDQOGLITLGMITHPTQAFLLSSVDLHTHCYQMMLP 360
Db 301 IPKQAGSDYCNTENEEELFLIDQOGLITLGMITHPTQAFLLSSVDLHTHCYQMMLP 360
Qy 361 ESVAIVCSPKFOETGFEKLTIDHGLEIBISSCRQGFHPSKDPFLFCSCHVTVDRAVTI 420
Db 361 ESVAIVCSPKFOETGFEKLTIDHGLEIBISSCRQGFHPSKDPFLFCSCHVTVDRAVTI 420
Qy 421 TDLR 424
Db 421 TDLR 424

RESULT 8
US-60-452-680-24059
Sequence 24059, Application US/60452680
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele
APPLICANT: GRUBE, Andrew
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CLO01450
CURRENT APPLICATION NUMBER: US/60/452,680
CURRENT FILING DATE: 2003-03-07
NUMBER OF SEQ ID NOS: 116213
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24059
LENGTH: 424
TYPE: PRT
ORGANISM: Homo sapiens

US-60-452-680-24059

Query Match 100.0%; Score 2208; DB 33; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3.1e-172;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMASTYSEGNIEHAF 60
 DB 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMASTYSEGNIEHAF 60
 QY 61 ILNNKYITLFIKLPKRDYKSAVPEKKDVKKLEIAFPKAEELKAEELKRYTKEYTE 120
 DB 61 ILNNKYITLFIKLPKRDYKSAVPEKKDVKKLEIAFPKAEELKAEELKRYTKEYTE 120
 QY 121 YNEEKKKEAEELARMAIIOELEKEKORVAQOQOQLSQEQFHAFEEMIRNOLEKERLK 180
 DB 121 YNEEKKKEAEELARMAIIOELEKEKORVAQOQOQLSQEQFHAFEEMIRNOLEKERLK 180
 QY 181 IVQFEGKVDPLGGLVVDLEKPSLDVFPPTLVSSIQPSDCHTTVPKPPVVDRLSKPG 240
 DB 181 IVQFEGKVDPLGGLVVDLEKPSLDVFPPTLVSSIQPSDCHTTVPKPPVVDRLSKPG 240
 QY 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMNEFTITHVL 300
 DB 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMNEFTITHVL 300
 QY 301 IPKOSAGSDYCNTENEEBELFLIQOQGLITLGMITHPTQAFSSVDLHTHCSYQMLP 360
 DB 301 IPKOSAGSDYCNTENEEBELFLIQOQGLITLGMITHPTQAFSSVDLHTHCSYQMLP 360
 QY 361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
 DB 361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
 QY 421 TDLR 424
 DB 421 TDLR 424

RESULT 9

US-60-452-680-24060
 ; Sequence 24060, Application US/60452680
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: CARGILL, Andrew
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CU001450
 ; CURRENT APPLICATION NUMBER: US/60/452, 680
 ; CURRENT FILING DATE: 2003-03-07
 ; NUMBER OF SEQ ID NOS: 116213
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24060
 ; LENGTH: 424
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-60-452-680-24060

Query Match 100.0%; Score 2208; DB 33; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3.1e-172;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMASTYSEGNIEHAF 60
 DB 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMASTYSEGNIEHAF 60
 QY 61 ILNNKYITLFIKLPKRDYKSAVPEKKDVKKLEIAFPKAEELKAEELKRYTKEYTE 120
 DB 61 ILNNKYITLFIKLPKRDYKSAVPEKKDVKKLEIAFPKAEELKAEELKRYTKEYTE 120
 QY 121 YNEEKKKEAEELARMAIIOELEKEKORVAQOQOQLSQEQFHAFEEMIRNOLEKERLK 180
 DB 121 YNEEKKKEAEELARMAIIOELEKEKORVAQOQOQLSQEQFHAFEEMIRNOLEKERLK 180

QY 181 IVQFEGKVDPLGGLVVDLEKPSLDVFPPTLVSSIQPSDCHTTVPKPPVVDRLSKPG 240
 DB 181 IVQFEGKVDPLGGLVVDLEKPSLDVFPPTLVSSIQPSDCHTTVPKPPVVDRLSKPG 240
 QY 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMNEFTITHVL 300
 DB 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMNEFTITHVL 300
 QY 301 IPKOSAGSDYCNTENEEBELFLIQOQGLITLGMITHPTQAFSSVDLHTHCSYQMLP 360
 DB 301 IPKOSAGSDYCNTENEEBELFLIQOQGLITLGMITHPTQAFSSVDLHTHCSYQMLP 360
 QY 361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
 DB 361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
 QY 421 TDLR 424
 DB 421 TDLR 424

RESULT 10

US-60-453-050-15016
 ; Sequence 15016, Application US/60453050
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: LIKE, May
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CU001457
 ; CURRENT APPLICATION NUMBER: US/60/453, 050
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15016
 ; LENGTH: 424
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-60-453-050-15016

Query Match 100.0%; Score 2208; DB 33; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3.1e-172;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMASTYSEGNIEHAF 60
 DB 1 MSDHGDVSLPPEDRVRLALSQLGSAVEVNEDIIPRRYFRSGVEIIRMASTYSEGNIEHAF 60
 QY 61 ILNNKYITLFIKLPKRDYKSAVPEKKDVKKLEIAFPKAEELKAEELKRYTKEYTE 120
 DB 61 ILNNKYITLFIKLPKRDYKSAVPEKKDVKKLEIAFPKAEELKAEELKRYTKEYTE 120
 QY 121 YNEEKKKEAEELARMAIIOELEKEKORVAQOQOQLSQEQFHAFEEMIRNOLEKERLK 180
 DB 121 YNEEKKKEAEELARMAIIOELEKEKORVAQOQOQLSQEQFHAFEEMIRNOLEKERLK 180
 QY 181 IVQFEGKVDPLGGLVVDLEKPSLDVFPPTLVSSIQPSDCHTTVPKPPVVDRLSKPG 240
 DB 181 IVQFEGKVDPLGGLVVDLEKPSLDVFPPTLVSSIQPSDCHTTVPKPPVVDRLSKPG 240
 QY 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMNEFTITHVL 300
 DB 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMNEFTITHVL 300
 QY 301 IPKOSAGSDYCNTENEEBELFLIQOQGLITLGMITHPTQAFSSVDLHTHCSYQMLP 360
 DB 301 IPKOSAGSDYCNTENEEBELFLIQOQGLITLGMITHPTQAFSSVDLHTHCSYQMLP 360
 QY 361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
 DB 361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420

QY 421 TDLR 424
 Db 421 TDLR 424

RESULT 11

US-60-453-050-15017
 ; Sequence 15017, Application US/60453050
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: LUKE, May
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C0001457
 ; CURRENT APPLICATION NUMBER: US/60/453,050
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15017
 ; LENGTH: 424
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-453-050-15017

Query Match 100.0%; Score 2208; DB 33; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3,1e-172;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGVSLPDPDRVRAISQSGAVENVEDIPPRYFRSGVEITIRMAISYSEGNIEHAF 60
 Db 1 MSDHGVSLPDPDRVRAISQSGAVENVEDIPPRYFRSGVEITIRMAISYSEGNIEHAF 60
 QY 61 ILYNKYITLPIEKLPRKRDYKSAVPEKKDTVKKLKEIAFPKAEELKRYTKYTE 120
 Db 61 ILYNKYITLPIEKLPRKRDYKSAVPEKKDTVKKLKEIAFPKAEELKRYTKYTE 120
 QY 121 YNEBKKEAEELARNMAIQOELEKORVAQOQOQOEJOFHAFEMIRNOSELERLK 180
 Db 121 YNEBKKEAEELARNMAIQOELEKORVAQOQOQOEJOFHAFEMIRNOSELERLK 180
 QY 181 IVOEFKGVDPGLGSPVLDLEKPSLDVPTLVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
 Db 181 IVOEFKGVDPGLGSPVLDLEKPSLDVPTLVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
 QY 241 ALSNSESPTIDGLRHVVVPGRLCPQLQASANTAGVTCGILGCKMRNEFTITHVL 300
 Db 241 ALSNSESPTIDGLRHVVVPGRLCPQLQASANTAGVTCGILGCKMRNEFTITHVL 300
 QY 301 IPKOSAGSDYCNTENEBELFLIDQOGLITLGMITHPTQTAFSLSVDLTHSCSYOMLP 360
 Db 301 IPKOSAGSDYCNTENEBELFLIDQOGLITLGMITHPTQTAFSLSVDLTHSCSYOMLP 360
 QY 361 ESWAIVCSPKQFQEPFLTDHGLEISSCRQGFPHSKDPLFCSCSHVTVDAVVI 420
 Db 361 ESWAIVCSPKQFQEPFLTDHGLEISSCRQGFPHSKDPLFCSCSHVTVDAVVI 420
 QY 421 TDLR 424
 Db 421 TDLR 424

RESULT 12

US-60-453-135-15016
 ; Sequence 15016, Application US/60453135
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: IAKOUBOVA, Olga
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C0001456
 ; CURRENT APPLICATION NUMBER: US/60/453,135
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 82762

; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15016
 ; LENGTH: 424
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-453-135-15016

Query Match 100.0%; Score 2208; DB 33; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3,1e-172;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGVSLPDPDRVRAISQSGAVENVEDIPPRYFRSGVEITIRMAISYSEGNIEHAF 60
 Db 1 MSDHGVSLPDPDRVRAISQSGAVENVEDIPPRYFRSGVEITIRMAISYSEGNIEHAF 60
 QY 61 ILYNKYITLPIEKLPRKRDYKSAVPEKKDTVKKLKEIAFPKAEELKRYTKYTE 120
 Db 61 ILYNKYITLPIEKLPRKRDYKSAVPEKKDTVKKLKEIAFPKAEELKRYTKYTE 120
 QY 121 YNEBKKEAEELARNMAIQOELEKORVAQOQOQOEJOFHAFEMIRNOSELERLK 180
 Db 121 YNEBKKEAEELARNMAIQOELEKORVAQOQOQOEJOFHAFEMIRNOSELERLK 180
 QY 181 IVOEFKGVDPGLGSPVLDLEKPSLDVPTLVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
 Db 181 IVOEFKGVDPGLGSPVLDLEKPSLDVPTLVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
 QY 241 ALSNSESPTIDGLRHVVVPGRLCPQLQASANTAGVTCGILGCKMRNEFTITHVL 300
 Db 241 ALSNSESPTIDGLRHVVVPGRLCPQLQASANTAGVTCGILGCKMRNEFTITHVL 300
 QY 301 IPKOSAGSDYCNTENEBELFLIDQOGLITLGMITHPTQTAFSLSVDLTHSCSYOMLP 360
 Db 301 IPKOSAGSDYCNTENEBELFLIDQOGLITLGMITHPTQTAFSLSVDLTHSCSYOMLP 360
 QY 361 ESWAIVCSPKQFQEPFLTDHGLEISSCRQGFPHSKDPLFCSCSHVTVDAVVI 420
 Db 361 ESWAIVCSPKQFQEPFLTDHGLEISSCRQGFPHSKDPLFCSCSHVTVDAVVI 420
 QY 421 TDLR 424
 Db 421 TDLR 424

RESULT 13

US-60-453-135-15017
 ; Sequence 15017, Application US/60453135
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele
 ; APPLICANT: IAKOUBOVA, Olga
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C0001456
 ; CURRENT APPLICATION NUMBER: US/60/453,135
 ; NUMBER OF SEQ ID NOS: 82762
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15017
 ; LENGTH: 424
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-453-135-15017

Query Match 100.0%; Score 2208; DB 33; Length 424;
 Best Local Similarity 100.0%; Pred. No. 3,1e-172;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGVSLPDPDRVRAISQSGAVENVEDIPPRYFRSGVEITIRMAISYSEGNIEHAF 60
 Db 1 MSDHGVSLPDPDRVRAISQSGAVENVEDIPPRYFRSGVEITIRMAISYSEGNIEHAF 60
 QY 61 ILYNKYITLPIEKLPRKRDYKSAVPEKKDTVKKLKEIAFPKAEELKRYTKYTE 120
 Db 61 ILYNKYITLPIEKLPRKRDYKSAVPEKKDTVKKLKEIAFPKAEELKRYTKYTE 120

```

Db      61 ILNNKYITLFIKLPKRDYKSAVPEKKDVTYKUKLKEIAPFAKABELKAKELKRYTKEYTE 120
Qy      121 YNEEKKEAEELARNMALIQOLEKEKORVAQOQOQLEQOFHAFEMIRNOLEKERLK 180
Db      121 YNEEKKEAEELARNMALIQOLEKEKORVAQOQOQLEQOFHAFEMIRNOLEKERLK 180
Qy      181 IVOEFGKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
Db      181 IVOEFGKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
Qy      241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMNEFTITHYL 300
Db      241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMNEFTITHYL 300
Qy      301 IPKOSAGSDYCNTENEEELFLIQOQGLITLGMHTHTPTQTAFSSVDLHTHCYQMLP 360
Db      301 IPKOSAGSDYCNTENEEELFLIQOQGLITLGMHTHTPTQTAFSSVDLHTHCYQMLP 360
Qy      361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Db      361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Qy      421 TDLR 424
Db      421 TDLR 424

```

RESULT 14

```

US-60-455-444-8192
; Sequence 8192, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; NUMBER OF SEQ ID NOS: 2003-03-18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8192
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-8192

```

```

Query Match      100.0%; Score 2208; DB 33; Length 424;
Best Local Similarity 100.0%; Pred. No. 3, 1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MSMDHGVSLPPEDRVRAISQLGSAVEVNEIIPRRYFRSGVEIIRMASIYSEGNIEHAF 60
Db      1 MSMDHGVSLPPEDRVRAISQLGSAVEVNEIIPRRYFRSGVEIIRMASIYSEGNIEHAF 60
Qy      61 ILNNKYITLFIKLPKRDYKSAVPEKKDVTYKUKLKEIAPFAKABELKAKELKRYTKEYTE 120
Db      61 ILNNKYITLFIKLPKRDYKSAVPEKKDVTYKUKLKEIAPFAKABELKAKELKRYTKEYTE 120
Qy      121 YNEEKKEAEELARNMALIQOLEKEKORVAQOQOQLEQOFHAFEMIRNOLEKERLK 180
Db      121 YNEEKKEAEELARNMALIQOLEKEKORVAQOQOQLEQOFHAFEMIRNOLEKERLK 180
Qy      181 IVOEFGKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
Db      181 IVOEFGKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
Qy      241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMNEFTITHYL 300
Db      241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMNEFTITHYL 300
Qy      301 IPKOSAGSDYCNTENEEELFLIQOQGLITLGMHTHTPTQTAFSSVDLHTHCYQMLP 360
Db      301 IPKOSAGSDYCNTENEEELFLIQOQGLITLGMHTHTPTQTAFSSVDLHTHCYQMLP 360

```

```

Qy      361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Db      361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Qy      421 TDLR 424
Db      421 TDLR 424

```

RESULT 15

```

US-60-455-444-8193
; Sequence 8193, Application US/60455444
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001455
; CURRENT APPLICATION NUMBER: US/60/455,444
; NUMBER OF SEQ ID NOS: 2003-03-18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8193
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-455-444-8193

```

```

Query Match      100.0%; Score 2208; DB 33; Length 424;
Best Local Similarity 100.0%; Pred. No. 3, 1e-172;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MSMDHGVSLPPEDRVRAISQLGSAVEVNEIIPRRYFRSGVEIIRMASIYSEGNIEHAF 60
Db      1 MSMDHGVSLPPEDRVRAISQLGSAVEVNEIIPRRYFRSGVEIIRMASIYSEGNIEHAF 60
Qy      61 ILNNKYITLFIKLPKRDYKSAVPEKKDVTYKUKLKEIAPFAKABELKAKELKRYTKEYTE 120
Db      61 ILNNKYITLFIKLPKRDYKSAVPEKKDVTYKUKLKEIAPFAKABELKAKELKRYTKEYTE 120
Qy      121 YNEEKKEAEELARNMALIQOLEKEKORVAQOQOQLEQOFHAFEMIRNOLEKERLK 180
Db      121 YNEEKKEAEELARNMALIQOLEKEKORVAQOQOQLEQOFHAFEMIRNOLEKERLK 180
Qy      181 IVOEFGKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
Db      181 IVOEFGKVDPLGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
Qy      241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMNEFTITHYL 300
Db      241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILCGKLMNEFTITHYL 300
Qy      301 IPKOSAGSDYCNTENEEELFLIQOQGLITLGMHTHTPTQTAFSSVDLHTHCYQMLP 360
Db      301 IPKOSAGSDYCNTENEEELFLIQOQGLITLGMHTHTPTQTAFSSVDLHTHCYQMLP 360
Qy      361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Db      361 ESVAIVCSPKFOETGFPLTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVVDRAVTI 420
Qy      421 TDLR 424
Db      421 TDLR 424

```

Search completed: August 17, 2004, 15:03:03
 Job time : 178 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 14:59:05 ; Search time 47 Seconds

(without alignments)
2832.024 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208

Sequence: 1 MSDHGVSLPDRVRALSQL.....FCSCSHVTVDRAVTTIDLR 424

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1292805 segs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	Published Applications AA:*
1:	/cgn2_6/ptodara/1/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodara/1/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodara/1/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodara/1/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodara/1/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodara/1/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodara/1/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodara/1/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodara/1/pubpaa/US09_PUBCOMB.pep:*
10:	/cgn2_6/ptodara/1/pubpaa/US09_PUBCOMB.pep:*
11:	/cgn2_6/ptodara/1/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodara/1/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodara/1/pubpaa/US10_PUBCOMB.pep:*
14:	/cgn2_6/ptodara/1/pubpaa/US10C_PUBCOMB.pep:*
15:	/cgn2_6/ptodara/1/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodara/1/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/ptodara/1/pubpaa/US60_PUBCOMB.pep:*
18:	/cgn2_6/ptodara/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	424	US-10-046-961-7	Sequence 7, Appl1
2	2208	100.0	424	US-10-340-578-66	Sequence 66, Appl1
3	2208	100.0	424	US-10-047-253-7	Sequence 7, Appl1
4	2208	100.0	424	US-10-671-572A-1	Sequence 1, Appl1
5	1904	86.2	424	US-10-671-572A-3	Sequence 3, Appl1
6	1269.5	57.5	421	US-10-046-961-5	Sequence 5, Appl1
7	1269.5	57.5	421	US-10-340-578-67	Sequence 67, Appl1
8	1269.5	57.5	421	US-10-047-253-5	Sequence 5, Appl1
9	1269.5	55.7	461	US-10-046-961-6	Sequence 6, Appl1
10	1269.5	55.7	461	US-10-340-578-68	Sequence 68, Appl1
11	1269.5	55.7	461	US-10-047-253-6	Sequence 6, Appl1
12	1269.5	55.7	461	US-10-012-600B-167	Sequence 167, Appl1
13	774.5	35.1	270	US-10-104-047-2064	Sequence 2064, Appl1
14	557.5	25.2	519	US-10-424-599-230882	Sequence 230882, Appl1
15	555.5	25.2	504	US-10-424-599-166616	Sequence 166616, Appl1

16	542	24.5	499	12	US-10-424-599-227385	Sequence 227385, Appl1
17	533.5	24.2	537	12	US-10-425-114-43326	Sequence 43326, Appl1
18	509.5	23.1	637	12	US-10-437-963-166644	Sequence 166644, Appl1
19	472	21.4	374	12	US-10-424-599-230878	Sequence 230878, Appl1
20	443	20.1	88	14	US-10-340-578-26	Sequence 26, Appl1
21	433	19.6	100	14	US-10-263-828-78	Sequence 78, Appl1
22	428.5	19.4	488	12	US-10-424-599-227382	Sequence 227382, Appl1
23	372	16.8	88	14	US-10-340-578-25	Sequence 25, Appl1
24	370	16.8	202	12	US-10-424-599-265209	Sequence 265209, Appl1
25	350.5	15.9	314	12	US-10-425-114-4475	Sequence 4475, Appl1
26	317	14.4	108	14	US-10-108-659-4756	Sequence 4756, Appl1
27	311	14.1	285	14	US-10-340-578-27	Sequence 27, Appl1
28	276	12.5	25	16	US-10-437-963-204466	Sequence 204466, Appl1
29	131.5	6.0	1175	9	US-09-771-161A-224	Sequence 224, Appl1
30	131.5	6.0	1175	9	US-09-771-161A-225	Sequence 225, Appl1
31	131.5	6.0	1175	9	US-09-771-161A-226	Sequence 226, Appl1
32	131.5	6.0	1212	12	US-10-168-582-9	Sequence 9, Appl1
33	131.5	6.0	1212	14	US-10-247-671-157	Sequence 157, Appl1
34	129	5.8	880	15	US-10-369-493-21643	Sequence 21643, Appl1
35	128	5.8	2701	14	US-10-171-311-83	Sequence 83, Appl1
36	126.5	5.7	3394	16	US-10-408-765A-529	Sequence 529, Appl1
37	126	5.7	1855	14	US-10-177-293-315	Sequence 315, Appl1
38	125	5.7	1233	10	US-09-291-437-89	Sequence 89, Appl1
39	124.5	5.6	380	15	US-10-161-927-84	Sequence 84, Appl1
40	124.5	5.6	519	15	US-10-108-260A-4747	Sequence 4747, Appl1
41	124.5	5.6	590	14	US-10-211-060-3	Sequence 3, Appl1
42	124	5.6	158	12	US-10-424-599-230880	Sequence 230880, Appl1
43	123.5	5.6	1118	14	US-10-153-668-104	Sequence 104, Appl1
44	122.5	5.5	305	14	US-10-128-714-3492	Sequence 3492, Appl1
45	122.5	5.5	593	14	US-10-128-714-8303	Sequence 8303, Appl1

ALIGNMENTS

RESULT 1	US-10-046-961-7
Sequence 7, Application US/10046961	
Publication No. US20020156012A1	
GENERAL INFORMATION:	
APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY	
APPLICANT: LYAPINA, Svetlana	
APPLICANT: LYAPINA, Rati	
APPLICANT: DESHAIES, Raymond	
TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS	
FILE REFERENCE: CIT1510-3	
CURRENT APPLICATION NUMBER: US/10/046,961	
CURRENT FILING DATE: 2002-01-14	
PRIOR APPLICATION NUMBER: US 60/261,314	
PRIOR FILING DATE: 2001-01-12	
PRIOR APPLICATION NUMBER: US 60/322,322	
PRIOR FILING DATE: 2001-09-14	
PRIOR APPLICATION NUMBER: US 60/322,030	
PRIOR FILING DATE: 2001-09-14	
NUMBER OF SEQ ID NOS: 22	
SOFTWARE: PatentIn version 3.1	
SEQ ID NO 7	
LENGTH: 424	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-10-046-961-7	
Query Match	100.0%; Score 2208; DB 13; Length 424;
Best Local Similarity	100.0%; Pred. No. 3e-174;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSDHGVSLPDRVRALSQLGSAVEVNEIDIPRRYFSGVEIIRMAISYSEGNIEHAF 60
DB	1 MSDHGVSLPDRVRALSQLGSAVEVNEIDIPRRYFSGVEIIRMAISYSEGNIEHAF 60
QY	1 ILYNKYITLFTLKRKHNDYSASVPEKKDYKKLKEIAFPAAELKAKELLKRYKETE 120
DB	61 ILYNKYITLFTLKRKHNDYSASVPEKKDYKKLKEIAFPAAELKAKELLKRYKETE 120

QY 121 YNEKKKEAEELARNNAIQOLEKEKQVAAQKQOQLEQEQFHAFEMIRNOELEKERLK 180
DB 121 YNEKKKEAEELARNNAIQOLEKEKQVAAQKQOQLEQEQFHAFEMIRNOELEKERLK 180
QY 181 IVOEFGKVDPLGGPLVDPDLKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
DB 181 IVOEFGKVDPLGGPLVDPDLKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
QY 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMRNEFTITHVL 300
DB 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMRNEFTITHVL 300
QY 301 IPKQASGSDYCNTENEEELFLIQOQGLITLGMHTPTQAFSSVDLTHGCSYQWMLP 360
DB 301 IPKQASGSDYCNTENEEELFLIQOQGLITLGMHTPTQAFSSVDLTHGCSYQWMLP 360
QY 361 ESVAIVCSPKFOETGFFKLTDHGLBEISSCRQKGFHPSKDPFLFCSCSHVTVVDRAVTI 420
DB 361 ESVAIVCSPKFOETGFFKLTDHGLBEISSCRQKGFHPSKDPFLFCSCSHVTVVDRAVTI 420
QY 421 TDLR 424
DB 421 TDLR 424

RESULT 2

US-10-340-578-66
; Sequence 66, Application US/10340578
; Publication No. US20030153097A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: DESHAIES, Raymond J.
; APPLICANT: COPE, Gregory
; APPLICANT: VERMA, Rati
; APPLICANT: AMEROGIO, Xavier I.
; TITLE OF INVENTION: MODULATION OF COP9 SIGNALSOME ISOPEPTIDASE ACTIVITY
; FILE REFERENCE: CIT1590-1
; CURRENT APPLICATION NUMBER: US/10/340,578
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/355,334
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 10/047,253
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 10/046,961
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-340-578-66

Query Match 100.0%; Score 2208; DB 14; Length 424;
Best Local Similarity 100.0%; Pred. No. 3e-174;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDHGDVSLPPEDRVRAISQSGSAVEVNEIDIPRRYFRSGVEIIRMASIYSEEGNIBAF 60
DB 1 MSDHGDVSLPPEDRVRAISQSGSAVEVNEIDIPRRYFRSGVEIIRMASIYSEEGNIBAF 60
QY 61 ILNNKYITLFTLEKLPKHADYSAVPEKKDVTYKLLKEIAPKAEELKAEILKKRYTKEYTE 120
DB 61 ILNNKYITLFTLEKLPKHADYSAVPEKKDVTYKLLKEIAPKAEELKAEILKKRYTKEYTE 120
QY 121 YNEKKKEAEELARNNAIQOLEKEKQVAAQKQOQLEQEQFHAFEMIRNOELEKERLK 180

DB 121 YNEKKKEAEELARNNAIQOLEKEKQVAAQKQOQLEQEQFHAFEMIRNOELEKERLK 180
QY 181 IVOEFGKVDPLGGPLVDPDLKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
DB 181 IVOEFGKVDPLGGPLVDPDLKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
QY 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMRNEFTITHVL 300
DB 241 ALSNSESIPITIDGLRHVVVPGRLCPQFLQASANTARGVETCGILCGKLMRNEFTITHVL 300
QY 301 IPKQASGSDYCNTENEEELFLIQOQGLITLGMHTPTQAFSSVDLTHGCSYQWMLP 360
DB 301 IPKQASGSDYCNTENEEELFLIQOQGLITLGMHTPTQAFSSVDLTHGCSYQWMLP 360
QY 361 ESVAIVCSPKFOETGFFKLTDHGLBEISSCRQKGFHPSKDPFLFCSCSHVTVVDRAVTI 420
DB 361 ESVAIVCSPKFOETGFFKLTDHGLBEISSCRQKGFHPSKDPFLFCSCSHVTVVDRAVTI 420
QY 421 TDLR 424
DB 421 TDLR 424

RESULT 3

US-10-047-253-7
; Sequence 7, Application US/10047253
; Publication No. US20030166243A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: COPE, Gregory
; APPLICANT: VERMA, Rati
; APPLICANT: ARAVIND, L
; APPLICANT: KOONIN, Eugene
; APPLICANT: DESHAIES, Raymond
; TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS
; FILE REFERENCE: CIT1510-4
; CURRENT APPLICATION NUMBER: US/10/047,253
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-047-253-7

Query Match 100.0%; Score 2208; DB 14; Length 424;
Best Local Similarity 100.0%; Pred. No. 3e-174;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDHGDVSLPPEDRVRAISQSGSAVEVNEIDIPRRYFRSGVEIIRMASIYSEEGNIBAF 60
DB 1 MSDHGDVSLPPEDRVRAISQSGSAVEVNEIDIPRRYFRSGVEIIRMASIYSEEGNIBAF 60
QY 61 ILNNKYITLFTLEKLPKHADYSAVPEKKDVTYKLLKEIAPKAEELKAEILKKRYTKEYTE 120
DB 61 ILNNKYITLFTLEKLPKHADYSAVPEKKDVTYKLLKEIAPKAEELKAEILKKRYTKEYTE 120
QY 121 YNEKKKEAEELARNNAIQOLEKEKQVAAQKQOQLEQEQFHAFEMIRNOELEKERLK 180
DB 121 YNEKKKEAEELARNNAIQOLEKEKQVAAQKQOQLEQEQFHAFEMIRNOELEKERLK 180
QY 181 IVOEFGKVDPLGGPLVDPDLKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240
DB 181 IVOEFGKVDPLGGPLVDPDLKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKPG 240

```
QY 241 ALSNSESIPTIDGLRHVVVGRGLCPQLQASANTARGVETGILCGKLMRNEFTITTHVL 300
DB 241 |||||
QY 301 IPKOSAGSDYCNTENEEBELFLIDQOGLITLGMHTHTPQTATFLSSVDLHTHCSYQMLP 360
DB 301 IPKOSAGSDYCNTENEEBELFLIDQOGLITLGMHTHTPQTATFLSSVDLHTHCSYQMLP 360
QY 361 ESWAIVCSPKFOETGFELTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
DB 361 ESWAIVCSPKFOETGFELTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
QY 421 TDLR 424
DB 421 TDLR 424

RESULT 4
US-10-671-572A-1
; Sequence 1, Application US/10671572A
; Publication No. US20040096954A1
; GENERAL INFORMATION:
; APPLICANT: SUGAMURA, Kazuo
; APPLICANT: TANAKA, Nobuyuki
; TITLE OF INVENTION: Protein AM5H and cDNA thereof
; FILE REFERENCE: 2003-1382/MMC/00653
; CURRENT APPLICATION NUMBER: US/10/671,572A
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: 09/831,452
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/JP99/06309
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: JP No. 10-322674
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 424
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-671-572A-1
```

```
Query Match 100.0%; Score 2208; DB 16; Length 424;
Best Local Similarity 100.0%; Pred. No. 3e-174;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSHDGVSLPPEDRVRLSQLGSAVEVNEIDIPRRYFRSGVEIIRMAISYEENIEHAF 60
DB 1 MSHDGVSLPPEDRVRLSQLGSAVEVNEIDIPRRYFRSGVEIIRMAISYEENIEHAF 60
QY 61 ILNKKYITLFIKLPKHRDYSAVPEKQDVKKLEIAFPKAEELKAEILKRYTKEYTE 120
DB 61 ILNKKYITLFIKLPKHRDYSAVPEKQDVKKLEIAFPKAEELKAEILKRYTKEYTE 120
QY 121 YNEBKKEAEELARNNAIQOELKEKORVAQOKOOLEBOFHAFEMIRNOLEKERLK 180
DB 121 YNEBKKEAEELARNNAIQOELKEKORVAQOKOOLEBOFHAFEMIRNOLEKERLK 180
QY 181 IVOEFGKVDPGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKRG 240
DB 181 IVOEFGKVDPGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKRG 240
QY 241 ALSNSESIPTIDGLRHVVVGRGLCPQLQASANTARGVETGILCGKLMRNEFTITTHVL 300
DB 241 ALSNSESIPTIDGLRHVVVGRGLCPQLQASANTARGVETGILCGKLMRNEFTITTHVL 300
QY 301 IPKOSAGSDYCNTENEEBELFLIDQOGLITLGMHTHTPQTATFLSSVDLHTHCSYQMLP 360
DB 301 IPKOSAGSDYCNTENEEBELFLIDQOGLITLGMHTHTPQTATFLSSVDLHTHCSYQMLP 360
QY 361 ESWAIVCSPKFOETGFELTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
DB 361 ESWAIVCSPKFOETGFELTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
```

```
QY 421 TDLR 424
DB 421 TDLR 424
```

```
RESULT 5
US-10-671-572A-3
; Sequence 3, Application US/10671572A
; Publication No. US20040096954A1
; GENERAL INFORMATION:
; APPLICANT: SUGAMURA, Kazuo
; APPLICANT: TANAKA, Nobuyuki
; TITLE OF INVENTION: Protein AM5H and cDNA thereof
; FILE REFERENCE: 2003-1382/MMC/00653
; CURRENT APPLICATION NUMBER: US/10/671,572A
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: 09/831,452
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: PCT/JP99/06309
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: JP No. 10-322674
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 424
; TYPE: PR
; ORGANISM: mouse
US-10-671-572A-3
```

```
Query Match 86.2%; Score 1904; DB 16; Length 424;
Best Local Similarity 83.5%; Pred. No. 4.7e-149;
Matches 354; Conservative 35; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 1 MSHDGVSLPPEDRVRLSQLGSAVEVNEIDIPRRYFRSGVEIIRMAISYEENIEHAF 60
DB 1 MSHDGVSLPPEDRVRLSQLGSAVEVNEIDIPRRYFRSGVEIIRMAISYEENIEHAF 60
QY 61 ILNKKYITLFIKLPKHRDYSAVPEKQDVKKLEIAFPKAEELKAEILKRYTKEYTE 120
DB 61 ILNKKYITLFIKLPKHRDYSAVPEKQDVKKLEIAFPKAEELKAEILKRYTKEYTE 120
QY 121 YNEBKKEAEELARNNAIQOELKEKORVAQOKOOLEBOFHAFEMIRNOLEKERLK 180
DB 121 YNEBKKEAEELARNNAIQOELKEKORVAQOKOOLEBOFHAFEMIRNOLEKERLK 180
QY 181 IVOEFGKVDPGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKRG 240
DB 181 IVOEFGKVDPGGLVLDLEKPSLDVFPPLTVSSIQPSDCHTTVPAPKPPVDRSLKRG 240
QY 241 ALSNSESIPTIDGLRHVVVGRGLCPQLQASANTARGVETGILCGKLMRNEFTITTHVL 300
DB 241 ALSNSESIPTIDGLRHVVVGRGLCPQLQASANTARGVETGILCGKLMRNEFTITTHVL 300
QY 301 IPKOSAGSDYCNTENEEBELFLIDQOGLITLGMHTHTPQTATFLSSVDLHTHCSYQMLP 360
DB 301 IPKOSAGSDYCNTENEEBELFLIDQOGLITLGMHTHTPQTATFLSSVDLHTHCSYQMLP 360
QY 361 ESWAIVCSPKFOETGFELTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
DB 361 ESWAIVCSPKFOETGFELTDHGLEIISCRQKGFHPSKDPPLFCSCSHVTVDRAVTI 420
QY 421 TDLR 424
DB 421 TDLR 424
```

```
RESULT 6
US-10-046-961-5
; Sequence 5, Application US/10046961
; Publication No. US20020156012A1
; GENERAL INFORMATION:
```

```
/ APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
/ APPLICANT: LYAPINIA, Svecliana
/ APPLICANT: VERMA, Rati
/ APPLICANT: DESHAIES, Raymond
/ TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS
/ FILE REFERENCE: CIT1510-3
/ CURRENT APPLICATION NUMBER: US/10/046,961
/ PRIOR FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US 60/261,314
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: US 60/322,322
/ PRIOR FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: US 60/322,030
/ PRIOR FILING DATE: 2001-09-14
/ NUMBER OF SEQ ID NOS: 22
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 421
/ TYPE: PR1
/ ORGANISM: Homo sapiens
US-10-046-961-5

Query Match      57.5%; Score 1269.5; DB 13; Length 421;
Best Local Similarity 55.8%; Pred. No. 1.7e-96;
Matches 240; Conservative 74; Mismatches 101; Indels 15; Gaps 2;

QY      1 MSDHGVSLPPEDEVRRLSQLSADEVNEDIPRRYFRSGVEITRMASITSEGNIEHAF 60
DB      1 MPHTDVSLSPEERVRRLSKLGCNITISDITPRYFRSGVEIMRNASVYLEGNLENAF 60
QY      61 ILVNYKITLFTIEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
DB      61 VLYNKFITLFEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
QY      61 VLYNKFITLFEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
DB      61 VLYNKFITLFEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
QY      121 YNEEKKEAEELANMAIQOLEKEKORVAQOQOOLEQOFHAFEMIRNOLEKERLK 180
DB      121 YLQSKNKYKAEILKLEHQRLEIAERKRIQMRQOQOQLESEQFLFFEDQLKQELARQMR 180
QY      121 YLQSKNKYKAEILKLEHQRLEIAERKRIQMRQOQOQLESEQFLFFEDQLKQELARQMR 180
DB      121 YLQSKNKYKAEILKLEHQRLEIAERKRIQMRQOQOQLESEQFLFFEDQLKQELARQMR 180
QY      181 IVQEFKVDGGLGGLPVDLEKPSLDVFPF-----LTVSSIQPSDCHTTPRAKPPVVD 234
DB      181 SQQTSG-----LSQIDGSLSCSTHNNNSLNVADQPNKSDATNVASHSPVN 231
QY      235 RSLKPGALNSSESIPITDGLRHVVVPGRLCPQFLQASANTARGETCGILCGKLMRNEF 294
DB      235 RSLKPGALNSSESIPITDGLRHVVVPGRLCPQFLQASANTARGETCGILCGKLMRNEF 294
QY      232 RALTPAATLSAVQNLVVEGLRCVLPEDLCHEFLQALAESNTVAGIRTCGILCGKLTNNEF 291
DB      232 RALTPAATLSAVQNLVVEGLRCVLPEDLCHEFLQALAESNTVAGIRTCGILCGKLTNNEF 291
QY      295 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 354
DB      295 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 354
QY      292 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 351
DB      292 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 351
QY      355 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 414
DB      355 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 414
QY      352 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 411
DB      352 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 411
QY      415 DRAVITIDLR 424
DB      415 DRAVITIDLR 424
QY      412 DIKIIVDLR 421
DB      412 DIKIIVDLR 421

RESULT 7
US-10-340-578-67
/ Sequence 67, Application US/10340578
/ Publication No. US20030153097A1
/ GENERAL INFORMATION:
/ APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
/ APPLICANT: DESHAIES, Raymond J.
/ APPLICANT: COPE, Gregory
/ APPLICANT: VERMA, Rati
/ APPLICANT: AMROGOLO, Xavier I.
/ TITLE OF INVENTION: MODULATION OF COP9 SIGNALSOME ISOPEPTIDASE ACTIVITY
/ FILE REFERENCE: CIT1590-1
/ CURRENT APPLICATION NUMBER: US/10/340,578
/ CURRENT FILING DATE: 2003-01-09
```

```
/ PRIOR APPLICATION NUMBER: US 60/355,334
/ PRIOR FILING DATE: 2002-02-06
/ PRIOR APPLICATION NUMBER: US 10/047,253
/ PRIOR FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US 10/046,961
/ PRIOR FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US 60/261,314
/ PRIOR FILING DATE: 2001-01-12
/ PRIOR APPLICATION NUMBER: US 60/322,322
/ PRIOR FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: US 60/322,030
/ PRIOR FILING DATE: 2001-09-14
/ NUMBER OF SEQ ID NOS: 83
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 67
/ LENGTH: 421
/ TYPE: PR1
/ ORGANISM: Homo sapiens
US-10-340-578-67

Query Match      57.5%; Score 1269.5; DB 14; Length 421;
Best Local Similarity 55.8%; Pred. No. 1.7e-96;
Matches 240; Conservative 74; Mismatches 101; Indels 15; Gaps 2;

QY      1 MSDHGVSLPPEDEVRRLSQLSADEVNEDIPRRYFRSGVEITRMASITSEGNIEHAF 60
DB      1 MPHTDVSLSPEERVRRLSKLGCNITISDITPRYFRSGVEIMRNASVYLEGNLENAF 60
QY      61 ILVNYKITLFTIEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
DB      61 VLYNKFITLFEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
QY      61 VLYNKFITLFEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
DB      61 VLYNKFITLFEKLPKRDYSAVTPKQDVKKLKEIAFPKAEILKAEILKRYTEKYTE 120
QY      121 YNEEKKEAEELANMAIQOLEKEKORVAQOQOOLEQOFHAFEMIRNOLEKERLK 180
DB      121 YLQSKNKYKAEILKLEHQRLEIAERKRIQMRQOQOQLESEQFLFFEDQLKQELARQMR 180
QY      121 YLQSKNKYKAEILKLEHQRLEIAERKRIQMRQOQOQLESEQFLFFEDQLKQELARQMR 180
DB      121 YLQSKNKYKAEILKLEHQRLEIAERKRIQMRQOQOQLESEQFLFFEDQLKQELARQMR 180
QY      181 IVQEFKVDGGLGGLPVDLEKPSLDVFPF-----LTVSSIQPSDCHTTPRAKPPVVD 234
DB      181 SQQTSG-----LSQIDGSLSCSTHNNNSLNVADQPNKSDATNVASHSPVN 231
QY      235 RSLKPGALNSSESIPITDGLRHVVVPGRLCPQFLQASANTARGETCGILCGKLMRNEF 294
DB      235 RSLKPGALNSSESIPITDGLRHVVVPGRLCPQFLQASANTARGETCGILCGKLMRNEF 294
QY      232 RALTPAATLSAVQNLVVEGLRCVLPEDLCHEFLQALAESNTVAGIRTCGILCGKLTNNEF 291
DB      232 RALTPAATLSAVQNLVVEGLRCVLPEDLCHEFLQALAESNTVAGIRTCGILCGKLTNNEF 291
QY      295 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 354
DB      295 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 354
QY      292 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 351
DB      292 TITHVILPKQSGAGSDYCNTENEBELFLIQOQGLITIGWHTHTPTOTAFISSVDLHTCS 351
QY      355 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 414
DB      355 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 414
QY      352 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 411
DB      352 YQMLPESVAIVCSPKQETGFFKLTGHEIISSCROKGFHPSKDPPLFCSCSHYTVV 411
QY      415 DRAVITIDLR 424
DB      415 DRAVITIDLR 424
QY      412 DIKIIVDLR 421
DB      412 DIKIIVDLR 421

RESULT 8
US-10-047-253-5
/ Sequence 5, Application US/10047253
/ Publication No. US20030166243A1
/ GENERAL INFORMATION:
/ APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
/ APPLICANT: COPE, Gregory
/ APPLICANT: VERMA, Rati
/ APPLICANT: ARAVIND, I.
/ APPLICANT: KOONIN, Eugene
/ APPLICANT: DESHAIES, Raymond
/ TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS
/ FILE REFERENCE: CIT1510-4
/ CURRENT APPLICATION NUMBER: US/10/047,253
/ CURRENT FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US 60/261,314
```

```

; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-047-253-5

```

```

Query Match          57.5%; Score 1269.5; DB 14; Length 421;
Best Local Similarity 55.8%; Pred. No. 1.7e-95;
Matches 240; Conservative 74; Mismatches 101; Indels 15; Gaps 2;

```

```

QY 1 MSDHGVSLPPEEDRVRLAQLQSAVEVNEDEPPRRYFRSGVEIIRMAISYSEGNIEHAF 60
Db 1 MPDHTDVSLSPPEERVALSKLGCNITISEDITPRRYFRSGVEMERMAASYLLEGNLENAF 60
QY 61 ILVNYKITLPIEKLPKPRDYKSAVIPKEDTVKLLKELAFPKAEELKAEILKRYKEYTE 120
Db 61 VLNYKFTLLVEKLPNHRDYQQCAVPEKODIMKKELIAPPRDELKNDLLKKYVEYQZ 120
QY 121 YNEKKKEAEELARNMAIQCELEKEKORVAQQOQLEQEPFAFEEMIRNOLEKERLK 180
Db 121 YLQSKNKYKAEILKLEHQRLIEARKRIAQWRQQQLSEQFLPFEDQIKKQELARGQWR 180
QY 181 IVQEGKVDPLGGLVPLDLEKPSLDVFP-----LTSSIQPSDCHTTAPAKPPVVD 234
Db 181 SQQTSG-----LSEQIDGSLSCFSTHQNNSILNVFADQPKSDATNVASHSPVN 231
QY 235 RSLKPGALNSSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGLMRNEF 294
Db 232 RALTPAATLSAVQNLVVEGLRCVLPEDLCHKFLQLABSNTVKGITCGILCGLTTHNEF 291
QY 295 TITHVLIIPKQASGSDYCNTENEBELFLIQDQGLITLGHHTHTPTQTAFLLSSVDLHTCS 354
Db 292 TITHVIVPKQASGDPYCDMENVEBELFNVQDQHDLLTLTGHTHTPTQAFLLSSVDLHTCS 351
QY 355 YQWMLPESSVAIVCSPKQGTGFPLTDHGLEBESSCQKGFPHSKDPPILFCGSHYTV 414
Db 352 YQMLPEALAIIVCSPKADGTGIFLITNAGMLEVSACKKGFPHHTKEPPLFSICKVLVK 411
QY 415 DRAVTITDLR 424
Db 412 DIKTIIVDLR 421

```

```

RESULT 9
US-10-046-961-6
; Sequence 6, Application US/10046961
; Publication No. US20020156012A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: LIAPINA, Svetlana
; APPLICANT: VERMA, Rati
; APPLICANT: DESHAIES, Raymond
; TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER PROTEINS
; FILE REFERENCE: CITI510-3
; CURRENT APPLICATION NUMBER: US/10/046,961
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 461

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-046-961-6
Query Match          55.7%; Score 1229.5; DB 13; Length 461;
Best Local Similarity 56.4%; Pred. No. 4e-93;
Matches 232; Conservative 72; Mismatches 92; Indels 15; Gaps 2;

```

```

QY 1 MSDHGVSLPPEEDRVRLAQLQSAVEVNEDEPPRRYFRSGVEIIRMAISYSEGNIEHAF 60
Db 16 MPDHTDVSLSPPEERVALSKLGCNITISEDITPRRYFRSGVEMERMAASYLLEGNLENAF 75
QY 61 ILVNYKITLPIEKLPKPRDYKSAVIPKEDTVKLLKELAFPKAEELKAEILKRYKEYTE 120
Db 76 VLNYKFTLLVEKLPNHRDYQQCAVPEKODIMKKELIAPPRDELKNDLLKKYVEYQZ 135
QY 121 YNEKKKEAEELARNMAIQCELEKEKORVAQQOQLEQEPFAFEEMIRNOLEKERLK 180
Db 136 YLQSKNKYKAEILKLEHQRLIEARKRIAQWRQQQLSEQFLPFEDQIKKQELARGQWR 195
QY 181 IVQEGKVDPLGGLVPLDLEKPSLDVFP-----LTSSIQPSDCHTTAPAKPPVVD 234
Db 196 SQQTSG-----LSEQIDGSLSCFSTHQNNSILNVFADQPKSDATNVASHSPVN 246
QY 235 RSLKPGALNSSESIPTIDGLRHVVVPGRLCPQFLQASANTARGVETGILCGLMRNEF 294
Db 247 RALTPAATLSAVQNLVVEGLRCVLPEDLCHKFLQLABSNTVKGITCGILCGLTTHNEF 306
QY 295 TITHVLIIPKQASGSDYCNTENEBELFLIQDQGLITLGHHTHTPTQTAFLLSSVDLHTCS 354
Db 307 TITHVIVPKQASGDPYCDMENVEBELFNVQDQHDLLTLTGHTHTPTQAFLLSSVDLHTCS 366
QY 355 YQWMLPESSVAIVCSPKQGTGFPLTDHGLEBESSCQKGFPHSKDPPILF 405
Db 367 YQMLPEALAIIVCSPKADGTGIFLITNAGMLEVSACKKGFPHHTKEPPLF 417

```

```

RESULT 10
US-10-340-578-68
; Sequence 68, Application US/10340578
; Publication No. US20030153097A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: DESHAIES, Raymond J.
; APPLICANT: COPE, Gregory
; APPLICANT: VERMA, Rati
; APPLICANT: AMEROGIO, Xavier I.
; TITLE OF INVENTION: MODULATION OF COP9 SIGNALSOME ISOPEPTIDASE ACTIVITY
; FILE REFERENCE: CITI590-1
; CURRENT APPLICATION NUMBER: US/10/340,578
; PRIOR FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/355,334
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 10/047,253
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 10/046,961
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/261,314
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/322,322
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/322,030
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-340-578-68

```

```

Query Match          55.7%; Score 1229.5; DB 14; Length 461;
Best Local Similarity 56.4%; Pred. No. 4e-93;

```


Page 7

```

QY      13  DRVAA-LSQGSAAVEVNEIDIPRRYRFRSGVEIIRMAISYEEGNIIEHAFILYKYLTLFI 71
Db      7  DRRNRINMAAAGVADVNRISLAFYRIRADNILRQADIFPAEKNIIDLYVMLRFSLSVS 66
QY      72  EKLPRKRDYKSAVPEKXDTYVK-----LKEIAFPKABELKAEILL-----KR 113
Db      67  ELLPRHDYSSP-PROKESIKTKLLISLNELENLKPVOQKINELNSKLAAYONGGKF 125
QY      114  YTKXYTYNEEKK-----KAEELARNMAIQOE 141
Db      126  SSNNSLDFSEYVKQTSASYLKAVRPTAGEFVYQGSRSQPSFVYRVEEHARLSLTLP 165
QY      142  LEXEK-----QVAAQKQOOLEQEQGFHAFEEKIRNOELEK- 176
Db      186  PPKKEETLSRHSILGPNGLKQGMWRPPIIDKIKYPSNIDLSFVELPSIQHSLDEBSJLKKD 245
QY      177  -----ERLKIYVQFGKVDPLGGPLVPLDEKPSLDVFLPVTSSIQPSDCHTVRP 227
Db      246  NSIAEHKKSLSLIDILIQSEDCQHPQOPQPHDDEPSSLIFETTERS---ALIEVIRQP 301
QY      228  AKPPVVD-RSLKPGA-----LSNSESIPITDGLAHVYVGR-----LCPQFLQASA 273
Db      302  SPPVLAEVQDLPAVSPCVNEAGCKTEIRSSSSVHVEAPMQLHISTALMESFPMKLAKS 361
QY      274  NTRAGVETGIILOGKLMRNEFTITHLIPQSGSPYCNENEEELFLIDOCGLITIGW 333
Db      362  NTKKNLTKCGLAGLLKNRKFYITALLIPQBSTSDSCOTTNEEELFEVQDKRSLFPLGW 421
QY      334  IHTPTQIAFLSSVDLHTHCSYOMLPESVAIYCSFK--FOETGFTKLT-DHGLBEISSC 390
Db      422  IHTPTQSCMSSIDLHTHSYQIMLPESVAIYMAPRDSRNRNGIFRLTRPGMSVTKOC 481
QY      391  RQKGFPHSKDP--PLFCGSHYTV-VDRAVITIDLR 424
Db      482  DQGFHHSQPPDGGPIYKTCIDVYNNPDLKFEVIDLR 519

RESULT 15
US-10-424-599-166616
? Sequence 166616, Application US/10424599
? Publication No. US20040031072A1
? GENERAL INFORMATION:
? APPLICANT: La Rosa Thomas J
? APPLICANT: Kovalic David K
? APPLICANT: Zhou Yihua
? APPLICANT: Cao Yongwei
? TITLE OR INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
? TITLE OR INVENTION: Plants and Uses Thereof for Plant Improvement
? FILE REFERENCES: 38-21(53223)B
? CURRENT APPLICATION NUMBER: US/10/424, 599
? CURRENT FILING DATE: 2003-04-28
? NUMBER OF SEQ. ID NOS: 285684
? SEQ. ID NO 166616
? LENGTH: 504
? TYPE: PRT
? ORGANISM: Glycine max
? FEATURE:
? OTHER INFORMATION: Clone ID: PAT_MRT3847_12146C.1 pep
US-10-424-599-166616

Query Match      25.2%; Score 555.5; DB 12; Length 504;
Best Local Similarity 30.28; Pred. No. 3.2e-37; Indels 109; Gaps 17;
Matches 153; Conservative 79; Mismatches 166;

QY      18  LSQGSAAVEVNEIDIPRRYFRSGVEIIRMAISYEEGNIIEHAFILYKYLTLFIKLPKH 77
Db      5  LDSIARVVEVDNRIIPRYRIRADNILLKQATIVREHNVVDLYIILLRFLSLSETLSYH 64
QY      78  RDKSAVPEPKDTYKKEIAFPKABELKAEILLKRYTYKXYTYNEEKKGE--ABELARN 135
Db      65  RDYQASIANERARN-KERSBAVIDELSELKPEFKLEVER--WNDSHVAKALLPEENGFN 119

```

```
QY 136 MAIQELEKEKORVACQKQ-----QQLEGE----- 160
Db 120 KALQSSVNSISIQWPAVNKQPAAGHSSQSSWNNYNNMLSDSWPIDKQFQKLSESLPPNKE 179
QY 161 --OFHAF--EMIRNQ-----ELEKERLKIYQEFKVD-----PGLGSP 195
Db 180 TLSHSHFLGNGLRGQWLGPSAIEIKVQYPSGSELTOAKDSSPNQAGLYDLVAIKDGDQSP 239
QY 196 LVPDLE-----KESLD--VFPTLVTSIQPSDCHTTVPAPKPPVDR----- 235
Db 240 VTSTMDSVLSLDGRWLRFPAVESCSPVYTESREDPLQLNLIKQPLPPPIAQYPERVPI 299
QY 236 -----SLKFGALSNSESIPTIDQL-----RHVVPGRLCPQFLQASANTAGVETCGI 284
Db 300 PPSKVADPRRGPAPAKSSH-----SGLGATTYQHLHIFVXWMEERLRLASENTRKNLETCV 355
QY 285 LCGKLMNEFTTHVLIPKQASAGSDYCNTENEEBLFIQDQGLITLGGWTHHTQTAFL 344
Db 356 LAGSLKRRVPHITLLIPKQESTSDSCOTLINEEIEFEVDSLSLFPJGWIHTHPSQTCFM 415
QY 345 SSVDLHTHCGYQWMLPESVAIVCSPKFQET--GFPKLTD--HGLEEISSCRQGFHPSK- 400
Db 416 SSVDLHTHYSQWMLPBAIIVMAFPTDTSPHGIFHLSDPGSVVIRNCOQRGFHHEEP 475
QY 401 --DPPLFCSCSHVTV--VDRAVTTIDLR 424
Db 476 EDGTPYEHCSHVYMNANLKFDVVDLR 502
```

Search completed: August 17, 2004, 15:04:41
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 14:56:39 ; Search time 20 Seconds
(without alignments)
1094.471 Million cell updates/sec

Title: US-09-831-452-1

Perfect score: 2208
Sequence: 1 MSDHGDVSLPPEDRVRLASQ.....FCCSCHVTVDRAVTTDLR 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A COMB.pep:.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep:.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep:.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep:.*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep:.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2208	100.0	424	US-09-665-479A-2	Sequence 2, Appl1
2	127.5	5.8	783	5231168-2	Patent No. 5231168
3	125	5.7	1233	US-09-688-188B-89	Sequence 89, Appl1
4	125	5.7	1233	US-09-291-417D-89	Sequence 89, Appl1
5	119	5.4	1239	US-09-688-188B-13	Sequence 13, Appl1
6	119	5.4	1239	US-09-291-417D-13	Sequence 13, Appl1
7	119	5.4	1239	US-09-418-710-70	Sequence 70, Appl1
8	118	5.3	1297	US-09-688-188B-14	Sequence 13, Appl1
9	118	5.3	1297	US-09-291-417D-14	Sequence 14, Appl1
10	118	5.3	1324	US-09-645-456A-13	Sequence 13, Appl1
11	118	5.3	1324	US-09-425-324A-13	Sequence 13, Appl1
12	118	5.3	1324	US-09-645-791-13	Sequence 13, Appl1
13	118	5.3	1332	US-09-645-456A-9	Sequence 9, Appl1
14	118	5.3	1332	US-09-425-324A-9	Sequence 9, Appl1
15	118	5.3	1332	US-09-425-324A-9	Sequence 9, Appl1
16	117	5.3	710	US-09-107-532A-5067	Sequence 5067, Appl1
17	117	5.3	1233	US-09-645-456A-35	Sequence 35, Appl1
18	117	5.3	1233	US-09-425-324A-35	Sequence 35, Appl1
19	117	5.3	1233	US-09-645-791-35	Sequence 35, Appl1
20	116.5	5.3	487	US-08-724-394A-7	Sequence 7, Appl1
21	115	5.3	1674	US-09-418-710-1	Sequence 1, Appl1
22	114	5.2	818	US-09-134-000C-6355	Sequence 6355, Appl1
23	114	5.2	2468	US-09-976-594-726	Sequence 726, Appl1
24	113	5.1	1805	US-07-853-703-2	Sequence 2, Appl1
25	111.5	5.0	3248	US-08-353-910-1	Sequence 1, Appl1
26	111.5	5.0	3248	PCT-US95-16216-1	Sequence 1, Appl1
27	111	5.0	1180	US-09-543-681A-6436	Sequence 6436, Appl1

28	110.5	5.0	376	6	5180810-1	Patent No. 5180810
29	110.5	5.0	1114	4	US-09-637-145-4	Sequence 4, Appl1
30	110	5.0	776	4	US-09-266-225D-10	Sequence 10, Appl1
31	110	5.0	1637	4	US-09-718-852-2	Sequence 2, Appl1
32	110	5.0	1637	4	US-09-718-852-2	Sequence 2, Appl1
33	110	5.0	1637	4	US-09-718-852-2	Sequence 2, Appl1
34	109.5	4.9	1244	4	US-09-418-710-21	Sequence 21, Appl1
35	109	4.9	1244	4	US-09-543-681A-6274	Sequence 6274, Appl1
36	109	4.9	1285	4	US-09-576-594-507	Sequence 507, Appl1
37	108.5	4.9	784	4	US-09-740-235-1	Sequence 1, Appl1
38	108.5	4.9	1353	4	US-09-645-456A-11	Sequence 11, Appl1
39	108.5	4.9	1353	4	US-09-425-324A-11	Sequence 11, Appl1
40	108.5	4.9	1353	4	US-09-645-791-11	Sequence 11, Appl1
41	108.5	4.9	1360	3	US-09-593-569-2	Sequence 2, Appl1
42	108.5	4.9	1360	4	US-09-579-664B-14	Sequence 14, Appl1
43	108.5	4.9	1360	4	US-09-645-456A-34	Sequence 34, Appl1
44	108.5	4.9	1360	4	US-09-425-324A-34	Sequence 34, Appl1
45	108.5	4.9	1360	4	US-09-645-791-34	Sequence 34, Appl1

ALIGNMENTS

```
RESULT 1
US-09-665-479A-2
Sequence 2, Application US/09665479A
Patent No. 6673570
GENERAL INFORMATION:
APPLICANT: Itch, Fumiko
APPLICANT: Itch, Susumu
APPLICANT: Heldin, Carl-Henrik
APPLICANT: ten-Dijke, Peter
TITLE OF INVENTION: SMAD ASSOCIATING POLYPEPTIDES
FILE REFERENCE: L00461.70096.US
CURRENT APPLICATION NUMBER: US/09/665,479A
CURRENT FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US 60/154,846
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 424
TYPE: PRT
ORGANISM: Homo sapiens
US-09-665-479A-2

Query Match      100.0%; Score 2208; DB 4; Length 424;
Best Local Similarity 100.0%; Pred. No. 6.1e-194;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDHGDVSLPPEDRVRLASQSAVEVNEIDIPPRYFSGVHTIMASTYSEGNIEHAF 60
DB 1 MSDHGDVSLPPEDRVRLASQSAVEVNEIDIPPRYFSGVHTIMASTYSEGNIEHAF 60
QY 61 ILYNXYITLFEKLPKHDYSAVPEKKDITVKLKEIAFPKAEELKRYKETE 120
DB 61 ILYNXYITLFEKLPKHDYSAVPEKKDITVKLKEIAFPKAEELKRYKETE 120
QY 61 ILYNXYITLFEKLPKHDYSAVPEKKDITVKLKEIAFPKAEELKRYKETE 120
DB 61 ILYNXYITLFEKLPKHDYSAVPEKKDITVKLKEIAFPKAEELKRYKETE 120
QY 121 YNEKKKAELARMAVIOQSLKEKORVAQKQQLFOEOPHAFEMIRNOELEKELK 180
DB 121 YNEKKKAELARMAVIOQSLKEKORVAQKQQLFOEOPHAFEMIRNOELEKELK 180
QY 121 YNEKKKAELARMAVIOQSLKEKORVAQKQQLFOEOPHAFEMIRNOELEKELK 180
DB 121 YNEKKKAELARMAVIOQSLKEKORVAQKQQLFOEOPHAFEMIRNOELEKELK 180
QY 181 IVQEGKXDPGAGLVPEDEKPSLDVPTLTSSIQSDCHTTPRPKPPVDSRLKPG 240
DB 181 IVQEGKXDPGAGLVPEDEKPSLDVPTLTSSIQSDCHTTPRPKPPVDSRLKPG 240
QY 181 IVQEGKXDPGAGLVPEDEKPSLDVPTLTSSIQSDCHTTPRPKPPVDSRLKPG 240
DB 181 IVQEGKXDPGAGLVPEDEKPSLDVPTLTSSIQSDCHTTPRPKPPVDSRLKPG 240
QY 241 AINSSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKXMEFTITVL 300
DB 241 AINSSESIPITDGLRHVVVPGRLCPQFLQASANTARGVETGILCGKXMEFTITVL 300
QY 301 IPKQASGSDYNTNEEELFLIOQOGLITGWIHTHTOTAFSSVVLTHGSGYQMLP 360
DB 301 IPKQASGSDYNTNEEELFLIOQOGLITGWIHTHTOTAFSSVVLTHGSGYQMLP 360
```

Query Match 5.8%; Score 127.5; DB 6; Length 783;
Best Local Similarity 21.2%; Pred. No. 0.0067;
Matches 62; Conservative 56; Mismatches 105; Indels 69; Gaps 14;

QY 361 ESVAIVCSPKFQETGFEKLTDLGLLEISSCQKGFPHSKDPLFCSCSHVTVVRAVVI 420
DB 361 ESVAIVCSPKFQETGFEKLTDLGLLEISSCQKGFPHSKDPLFCSCSHVTVVRAVVI 420

QY 421 TDLR 424
DB 421 TDLR 424

RESULT 2
5231168-2
; Patent No. 5231168
; APPLICANT: DZIEGIEL, MORTEN, BORRE, MARTIN, JENSEN, SOREN;
; VUSST, JENS, RIENECK, KLAUS, WIND, ANNETTE, JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 2
; LENGTH: 783

Query Match 5.8%; Score 127.5; DB 6; Length 783;
Best Local Similarity 21.2%; Pred. No. 0.0067;
Matches 62; Conservative 56; Mismatches 105; Indels 69; Gaps 14;

QY 9 LPEDRVRLSOLGSAVENEDIPRRYFRSGVEIIRMAISYEEN--IEHAFILYNYK 66
DB 429 LPEDDKNEKVEH--EIVEVEEILPEDDKNEKQHEIVEVEEILPEDDKNEKQHEIVEVE-- 484

QY 67 ITLFIETLPHKRDYK-----SAVPEKKD-----TVKKLEIAFPKAEELKA----- 108
DB 485 -----ELLPEDKNEKQHEIVEVEEILPEDDKNEKQHEIVEVEEILPEDDKNEKQHEIV 538

QY 109 ---ELTKRYTEYETNEEKKKAELLA---RNMAIQOLEKEKQKVAQKQOQLEQEF 162
DB 539 EVELILPEBEX--NEKQHEIVEVEEILPEDDKNEKQHEIVEVEEILPEDDKNEKQHEIV 596

QY 163 HAF-----EIRNOELEKELKIVQ-----EFGKDPGLGGLVVDLEKPSLDYF 208
DB 597 EVELILPEIVEIEEVPQINNENIETIKPEKKNEFS-----VEKAIPOBPV 646

QY 209 PTLTVS---SIOPSDCHTVPAKPPVDRSLKFGALSNSESIPTIDGLRHV 257
DB 647 PTLNENNVTPKPSGEGST-----KPDIVQIKIVQENKPKETPPVVDGPKHV 694

RESULT 3
US-09-688-188B-89
; Sequence 89, Application US/09688188B
; Parent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 1233
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-688-188B-89

Query Match 5.7%; Score 125; DB 4; Length 1233;
Best Local Similarity 21.9%; Pred. No. 0.023;
Matches 50; Conservative 50; Mismatches 76; Indels 52; Gaps 9;

QY 26 EVNEDIPPRRYFRS-----GVEIIRMAISYSEGNIEHAFILYNYKITLFIETLPHKRD 79
DB 327 EEESEVPQEGEPSSIVNVPGESTLRDFLRLOQENKRSALRQOL--IQEOQLREQEE 385

QY 80 YKSAVIEPKDVTYKLE--IAFPKAEELKAEELKRYTKETNEEKKKAELARMA 137
DB 386 YKQOLARQKRIQOQKORRRLEQORREHARQOQREORRRQOQREKRLBELERRK 445

QY 138 IQOE---LEKQKQVAQKQ---QOLEBOQFAFEEMIRNOELEKELKIVQEFQVDPG 191
DB 446 EEEERRRREKRRVREHQQEYIRQLBEQOHL--EILQOOLQEQAM----- 491

QY 192 LGGLVPLDLEKPSLDVFPPTLVSSIOPSDCHTVPAKPPVDRSLK 239
DB 492 ----LLHDDRPP-----HAQQQPPPPQOQDRS--KP 516

RESULT 4
US-09-291-417D-89
; Sequence 89, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 1233
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-291-417D-89

Query Match 5.7%; Score 125; DB 4; Length 1233;
Best Local Similarity 21.9%; Pred. No. 0.023;
Matches 50; Conservative 50; Mismatches 76; Indels 52; Gaps 9;

QY 26 EVNEDIPPRRYFRS-----GVEIIRMAISYSEGNIEHAFILYNYKITLFIETLPHKRD 79
DB 327 EEESEVPQEGEPSSIVNVPGESTLRDFLRLOQENKRSALRQOL--IQEOQLREQEE 385

QY 80 YKSAVIEPKDVTYKLE--IAFPKAEELKAEELKRYTKETNEEKKKAELARMA 137
DB 386 YKQOLARQKRIQOQKORRRLEQORREHARQOQREORRRQOQREKRLBELERRK 445

QY 138 IQOE---LEKQKQVAQKQ---QOLEBOQFAFEEMIRNOELEKELKIVQEFQVDPG 191
DB 446 EEEERRRREKRRVREHQQEYIRQLBEQOHL--EILQOOLQEQAM----- 491

QY 192 LGGLVPLDLEKPSLDVFPPTLVSSIOPSDCHTVPAKPPVDRSLK 239
DB 492 ----LLHDDRPP-----HAQQQPPPPQOQDRS--KP 516

RESULT 5
US-09-688-188B-13
; Sequence 13, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

FILE REFERENCE: 038602/0328
CURRENT APPLICATION NUMBER: US/09/688,188B
CURRENT FILING DATE: 2000-10-16
PRIOR APPLICATION NUMBER: 09/291,417
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: 60/081,784
PRIOR FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 155
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 13
LENGTH: 1239
TYPE: PRT
ORGANISM: Homo sapiens
US-09-688-188B-13

Query Match 5.4%; Score 119; DB 4; Length 1239;
Best Local Similarity 21.2%; Pred. No. 0.091;
Matches 65; Conservative 57; Mismatches 108; Indels 76; Gaps 12;

QY 26 EVNEDIPRRYFRS-----GVEIIRMASIYSEEGNIEHAFLIYNKYITLFIKLPKHRD 79
DB 328 EEEVEVEQEGEPSSIVNVGSESTLRDPLRLQENKERESEALRRQOL-LQEQQLREQOE 386
QY 80 YKSAVIEPKDVTVKLKE--IAPKABELKALIKRYTEYENEEKKEAEELARNA 137
DB 387 YKROLAERQKRIEQKEQRRRLLEEQRRERARQOEERQRRRDEKRLSELERRK 446
QY 138 IOOE--LEKEKQVAQOKQ---QOLBOQPH-----AFEEKTRNOE 173
DB 447 EEEERRRAEERKRVREGEYIRQLBEEQRHLEVLQOQLLOEQAMLLECRWREMEHRQ 506
QY 174 LEKERLKIYOE-----FGKYDPGLGGLVPLDLKPSL----- 205
DB 507 AERLQRLQOEQAVYLSLQDHRHRPHQSOQPP---PQOEBSKPSFHAEPPKAYEP 562
QY 206 -----DVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPGALSNSESI-PTIDGLRH 256
DB 563 ADRAAREVEDRFRKTNHSSPE-AQSKOTGRVLEPPVPSRS-ESFSNGSSESVHPLQRPAA 620
QY 257 VVYVGR 262
DB 621 PVPVPR 626

RESULT 6
US-09-291-417D-13
Sequence 13, Application US/09291417D
Patent No. 6680170
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHYTE, DAVID
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
FILE REFERENCE: 038602/0329
CURRENT APPLICATION NUMBER: US/09/291,417D
CURRENT FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/081,784
PRIOR FILING DATE: 1998-04-14
NUMBER OF SEQ ID NOS: 155
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 13
LENGTH: 1239
TYPE: PRT
ORGANISM: Homo sapiens
US-09-291-417D-13

Query Match 5.4%; Score 119; DB 4; Length 1239;
Best Local Similarity 21.2%; Pred. No. 0.081;
Matches 65; Conservative 57; Mismatches 108; Indels 76; Gaps 12;

QY 26 EVNEDIPRRYFRS-----GVEIIRMASIYSEEGNIEHAFLIYNKYITLFIKLPKHRD 79
DB 328 EEEVEVEQEGEPSSIVNVGSESTLRDPLRLQENKERESEALRRQOL-LQEQQLREQOE 386

QY 80 YKSAVIEPKDVTVKLKE--IAPKABELKALIKRYTEYENEEKKEAEELARNA 137
DB 387 YKROLAERQKRIEQKEQRRRLLEEQRRERARQOEERQRRRDEKRLSELERRK 446
QY 138 IOOE--LEKEKQVAQOKQ---QOLBOQPH-----AFEEKTRNOE 173
DB 447 EEEERRRAEERKRVREGEYIRQLBEEQRHLEVLQOQLLOEQAMLLECRWREMEHRQ 506
QY 174 LEKERLKIYOE-----FGKYDPGLGGLVPLDLKPSL----- 205
DB 507 AERLQRLQOEQAVYLSLQDHRHRPHQSOQPP---PQOEBSKPSFHAEPPKAYEP 562
QY 206 -----DVFPPLTVSSIQPSDCHTTPRAKPPVDRSLKPGALSNSESI-PTIDGLRH 256
DB 563 ADRAAREVEDRFRKTNHSSPE-AQSKOTGRVLEPPVPSRS-ESFSNGSSESVHPLQRPAA 620
QY 257 VVYVGR 262
DB 621 PVPVPR 626

RESULT 7
US-09-418-710-70
Sequence 70, Application US/09418710
Patent No. 6596482
GENERAL INFORMATION:
APPLICANT: Jones, Michael H.
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
FILE REFERENCE: 06501-042001
CURRENT APPLICATION NUMBER: US/09/418,710
CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/JP98/01783
PRIOR FILING DATE: 1998-04-17
PRIOR APPLICATION NUMBER: JP 9/310027
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: JP 9/116570
NUMBER OF SEQ ID NOS: 73
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 70
LENGTH: 1673
TYPE: PRT
ORGANISM: Homo sapiens
US-09-418-710-70

Query Match 5.4%; Score 119; DB 4; Length 1673;
Best Local Similarity 22.2%; Pred. No. 0.13;
Matches 114; Conservative 61; Mismatches 159; Indels 180; Gaps 27;

QY 21 LGSAAVENEDIPRRYFRS-----GVEIIRMASIY-----SEEGNIE 57
DB 675 LAGSADVTSANAKRYIQKGGCFDATTDAQMEELRLSNPBLVKLSTSYVDLTPEGKATIL 734
QY 58 HAFIILYNKYITL-----FIEKLPRHDYKSAVIEPKDVTVKLKEIAPKAEELKAEELK 112
DB 735 HA--LOGKLLTLVSTRDIE-----DYVD-ILRQAKQEPRLKAEQRRKEREBAALIR 785
QY 113 RYTKYETENEEKKEAEELAR-----NMAIQELEKQVAQOK---QOQLBOE 160
DB 786 KRKEKLEKEQEKMEKEKLEKEDQORNSSTADISGEBEREDFTSISKQTEQKELQOD 845
QY 161 QF-----HAFEEKTRNOE-----EKERLKIYQFGKVD 189
DB 846 MTEDEDDPGSHKRRRGKRGQNGFKFTROQINCVRRELLTAAEBEALK--QENQKRE 903
QY 190 PELGG-----PLVPDLKPSLDVFPPLTVSSIQ-----PSDCHT 223
DB 904 KELLEKLOSALACTNIFPLGRDMYRRYVIFPSLGLFIEEDYSGLTEDHLLPRSSFCN 963
QY 224 TVRAKPPVDRSLKPGALSNSESIPTID-GLRH-----VVYVGR-----LCPPQLQ 269
DB 964 NVQSQDPQV---STTGTPLMSSESTSNIDQGRDHSVQLPKRVHKNPKWCFYSSCEQLDQ 1020

RESULT 9
US-09-291-417D-14
Sequence 14, Application US/09291417D
Patient No. 6680170
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: MARTINEZ, RICARDO
APPLICANT: WHYTE, DAVID
TITLE OF INVENTION: STI20-RELATED PROTEIN KINASES
FILE REFERENCE: 038602/0329
CURRENT APPLICATION NUMBER: US/09/291,417D
CURRENT FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/081,784
PRIOR FILING DATE: 1998-04-14

```

Qy      76  KHRYSAAVLPKKDTQALKE--IAPPKAELTAEILTKYTKATYVNEKKAEELA 133
Db      379  ENEHKKQLLAERQKRIEEOKEQRRRLIEEOQRREKELRKQOEQRGRHHYEQMRREE-R 437

Qy      134  RNAAIOEELK--EKORVAQCKQOOLEQEFHAEEMINQF--LEKERL----- 179
Db      438  RRAHEQEYVRKQLIEQRQAERLQRLQKQRDYLVSIQHROEQRPVEKKPLVHYKEGMS 497

Qy      180  -----KIVQELGKVD---PQLGGPLVYDLEKESLDV-FPLTYVSSIQPSDCHTIV 226
Db      498  PSEKPAWAKEVEEERSLRNRQSSPMPHPKVNRIISDPLPRSEBSFISGVQFATPPMLR 557

Qy      227  PAKRP-----VYDRSLKPGALSNSESI---PT 250
Db      558  PVDDPQIPLVAVKSQGP-ALTAASQVAYEQFT 567

```

RESULT 11
US-09-425-324A-13
; Sequence 13, Application US/09425324A
; Patent No. 6562591
; GENERAL INFORMATION:
; APPLICANT: Fu, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/425,324A
; PRIORITY FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1324
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-425-324A-13

Query Match 5.3%; Score 118; DB 4; Length 1324;

Best Local Similarity 23.7%; Pred. No. 0.11; Indels 38; Gaps 10;
Matches 50; Conservative 50; Mismatches 73;

QY 76 KHRDYKSAVTPKEDYVKKLE--IAFPKAEELKAEELKRYKETEYNEKKKEAEELA 133
DB 379 ENEHKKQLLAERKRIEEOQERRLLEEOQRREKELRKQERQRHRYEEOQRREEE-R 437
QY 134 RNMAIQOELK--EKQRYAQOQOOLEOQFHAPEEMIRNOE--LEKRL----- 179
DB 438 RRAHEQGYKQLEOQOAEERLQRLKQERDYLVSLQHQRQORPVEKKPLHYHKEGMS 497
QY 180 -----KIVQEFKVD---PGLGGPLVPLEKPSLDV--PFTLVSSIQPSDCHTTR 226
DB 498 PSEKPAKAEVEERSRINQSSPAMPKVKANRISDPNLPRESSESISGVQAPARTPMLR 557
QY 227 PAKPP---VVDRLKPGALSNSESI---PT 250
DB 558 PVDPOIPLHLVAVKSQGP-ALTASQSVHEOPT 587

RESULT 12

US-09-645-791-13
; Sequence 13, Application US/09645791
; Patent No. 6569658
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Shen, Alan C
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
; FILE REFERENCE: A-68344-1/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/645,791
; PRIORITY FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 09/425,324
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1324
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-645-791-13

Query Match 5.3%; Score 118; DB 4; Length 1324;

Best Local Similarity 23.7%; Pred. No. 0.11; Indels 38; Gaps 10;
Matches 50; Conservative 50; Mismatches 73;

QY 76 KHRDYKSAVTPKEDYVKKLE--IAFPKAEELKAEELKRYKETEYNEKKKEAEELA 133
DB 379 ENEHKKQLLAERKRIEEOQERRLLEEOQRREKELRKQERQRHRYEEOQRREEE-R 437

DB 379 ENEHKKQLLAERKRIEEOQERRLLEEOQRREKELRKQERQRHRYEEOQRREEE-R 437
QY 134 RNMAIQOELK--EKQRYAQOQOOLEOQFHAPEEMIRNOE--LEKRL----- 179
DB 438 RRAHEQGYKQLEOQOAEERLQRLKQERDYLVSLQHQRQORPVEKKPLHYHKEGMS 497
QY 180 -----KIVQEFKVD---PGLGGPLVPLEKPSLDV--PFTLVSSIQPSDCHTTR 226
DB 498 PSEKPAKAEVEERSRINQSSPAMPKVKANRISDPNLPRESSESISGVQAPARTPMLR 557
QY 227 PAKPP---VVDRLKPGALSNSESI---PT 250
DB 558 PVDPOIPLHLVAVKSQGP-ALTASQSVHEOPT 587

RESULT 13

US-09-645-456A-9
; Sequence 9, Application US/09645456A
; Patent No. 6562580
; GENERAL INFORMATION:
; APPLICANT: Fu, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/645,456A
; PRIORITY FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US/09/425,324
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-645-456A-9

Query Match 5.3%; Score 118; DB 4; Length 1332;

Best Local Similarity 23.7%; Pred. No. 0.11; Indels 38; Gaps 10;
Matches 50; Conservative 50; Mismatches 73;

QY 76 KHRDYKSAVTPKEDYVKKLE--IAFPKAEELKAEELKRYKETEYNEKKKEAEELA 133
DB 379 ENEHKKQLLAERKRIEEOQERRLLEEOQRREKELRKQERQRHRYEEOQRREEE-R 437
QY 134 RNMAIQOELK--EKQRYAQOQOOLEOQFHAPEEMIRNOE--LEKRL----- 179
DB 438 RRAHEQGYKQLEOQOAEERLQRLKQERDYLVSLQHQRQORPVEKKPLHYHKEGMS 497
QY 180 -----KIVQEFKVD---PGLGGPLVPLEKPSLDV--PFTLVSSIQPSDCHTTR 226
DB 498 PSEKPAKAEVEERSRINQSSPAMPKVKANRISDPNLPRESSESISGVQAPARTPMLR 557
QY 227 PAKPP---VVDRLKPGALSNSESI---PT 250
DB 558 PVDPOIPLHLVAVKSQGP-ALTASQSVHEOPT 587

RESULT 14

US-09-425-324A-9
; Sequence 9, Application US/09425324A
; Patent No. 6562591
; GENERAL INFORMATION:
; APPLICANT: Fu, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/425,324A
; PRIORITY FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9


```
LENGTH: 1332
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-425-324A-9
```

```
Query Match          5.3%; Score 118; DB 4; Length 1332;
Best Local Similarity 23.7%; Pred. No. 0.11;
Matches 50; Conservative 50; Mismatches 73; Indels 38; Gaps 10;
```

```
QY 76 KHRDYKSAVPEKKDYKLUKE--IAPPKAEELKABLKRKYTEYBNEEKKKEABELA 133
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 379 ENEHKKQLAEKRIEEOQRRLIEOQRREKELKQEREQRHRYEQMRREE-R 437
QY 134 RNMAIOELEK---EKORVAQOKOOLEEOQFAFEEIRNOE---LEKERL----- 179
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 438 RRAHEOEYRKQLEORQAKRLQRLQKREDYVLSIQHORQORPEKKPLHYHKEGMS 497
QY 180 -----KIVQEFGRVD---PGLGGLVDPDLKPSLDV-FPTLVSSIQPSDCHTVR 226
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 498 PSEKPAWAKEVEERSRLNQSPFAMPKPVANRISDPYLPKPSSEFSISGVQPARTPMLR 557
QY 227 PAKPP---VYDRSLKPGALSNSESI---PT 250
Db 558 PVDQIPLHLVAVKSGP-ALTASQSVHEQPT 587
```

RESULT 15

```
US-09-645-791-9
Sequence 9, Application US/09645791
Patent No. 6569658
GENERAL INFORMATION:
APPLICANT: Luo, Ying
APPLICANT: Fu, Alan C
APPLICANT: Shen, Mary
TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS; COMPOSITIONS AND
TITLE OF INVENTION: METHODS OF USE
FILE REFERENCE: A-68344-1/RMS/DBR
CURRENT APPLICATION NUMBER: US/09/645,791
CURRENT FILING DATE: 2000-08-24
PRIORITY APPLICATION NUMBER: US 09/425,324
PRIORITY FILING DATE: 1999-10-21
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 1332
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: synthetic
US-09-645-791-9
```

```
Query Match          5.3%; Score 118; DB 4; Length 1332;
Best Local Similarity 23.7%; Pred. No. 0.11;
Matches 50; Conservative 50; Mismatches 73; Indels 38; Gaps 10;
```

```
QY 76 KHRDYKSAVPEKKDYKLUKE--IAPPKAEELKABLKRKYTEYBNEEKKKEABELA 133
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 379 ENEHKKQLAEKRIEEOQRRLIEOQRREKELKQEREQRHRYEQMRREE-R 437
QY 134 RNMAIOELEK---EKORVAQOKOOLEEOQFAFEEIRNOE---LEKERL----- 179
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 438 RRAHEOEYRKQLEORQAKRLQRLQKREDYVLSIQHORQORPEKKPLHYHKEGMS 497
QY 180 -----KIVQEFGRVD---PGLGGLVDPDLKPSLDV-FPTLVSSIQPSDCHTVR 226
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 498 PSEKPAWAKEVEERSRLNQSPFAMPKPVANRISDPYLPKPSSEFSISGVQPARTPMLR 557
QY 227 PAKPP---VYDRSLKPGALSNSESI---PT 250
Db 558 PVDQIPLHLVAVKSGP-ALTASQSVHEQPT 587
```